

# Bayesian Modelling of Pattern Formation from One Snapshot of Pattern

Natsuhiko Yoshinaga\*

*WPI - Advanced Institute for Materials Research,  
Tohoku University, Sendai 980-8577, Japan and  
MathAM-OIL, AIST, Sendai 980-8577, Japan*

Satoru Tokuda

*Research Institute for Information Technology, Kyushu University, Kasuga 816-8580, Japan and  
MathAM-OIL, AIST, Sendai 980-8577, Japan*

## CONTENTS

SI. Summary of this supplementary material	1
SII. Related works	2
A. Pattern formation and crystalline structure	2
B. Inverse structural optimisation	2
C. parameter estimation of governing equations	3
D. State-space model	3
1. Gaussian process	5
2. shooting method	5
3. filtering and sequential data assimilation	6
4. adjoint method and sensitivity analysis	6
E. estimation of the model with a stationary state	6
SIII. Family of models	7
A. Phase-field crystal model and its generalisation	7
B. Polynomial expansion of the linear operation with respect to wavenumbers	7
C. Expansion of the linear operator by orthogonal bases	8
D. design of nonlinear terms	9
SIV. Examples	11
A. Target pattern produced by a numerical simulation: two-dimensional dodecagonal quasi-crystal	11
B. Target pattern synthesised by a function: two-dimensional dodecagonal quasi-crystal	13
C. Target pattern synthesised by a function: stripe and hexagonal patterns	14
D. Target pattern synthesised by a function: double gyroid and Frank Kasper A15	18
SV. Damaged target pattern	22
References	23

## SI. SUMMARY OF THIS SUPPLEMENTARY MATERIAL

The purpose of this supplementary material is to show detailed data to support the main text. They are not necessary to understand our main claim, but we hope this supplementary material would be helpful for deeper understanding. Section SII is a detailed comparison between our method to other methods to estimate a governing equation from data. In Sec. SIII, we show the lengthy form of  $\mathcal{L}_k$  in equation (11) in Methods of the main text, and generalisation of the models. In Sec. SIV, the detailed data of the estimation is shown. We also demonstrate the estimation for the damaged target pattern in Sec. SV.

---

\* corresponding authors: yoshinaga@tohoku.ac.jp

## SII. RELATED WORKS

Our method covers interdisciplinary research fields such as materials science of soft materials and crystalline structures, nonlinear dynamics of pattern formation, and data-driven science techniques of machine learning and inverse problems. In this section, we review related works.

### A. Pattern formation and crystalline structure

First, we give a brief review of nonlinear dynamics on pattern formation and self-assembly of crystalline-like structures. The nonlinear partial differential equations such as the Swift-Hohenberg equation (SH) [97] have been used to describe a periodic pattern as a minimal model. SH was originally derived for stationary finite number instability of fluid convection[97–101], and then has later been applied to optics[102], ecosystem[103], and other seemingly distinct problems. The Ohta-Kawasaki-Oono (OKO) equation, which is similar to the SH equation, was proposed to describe the micro-phase separation of block copolymers[104, 105], whereas SH may also have been used for the same problem[106, 107]. In the theory of block copolymers, the SH equation is instead called the Landau-Brazovskii theory. These equations reproduce a stripe (also called lamellar or smectic) and hexagonal patterns in two dimensions[108], and lamellar, hexagonal cylinders, BCC, hexagonal closed packing patterns, also gyroid patterns in three dimensions [107, 109, 110]. The essence of these models is that there is a specific length scale, or wavenumber, at which the uniform state becomes unstable. Recently, the conserved version of SH has been used for the description of the crystalline structures [111, 112]. This model is specifically called the phase-field crystal model (PFC). The models that we used in this work are extensions of the PFC.

All these models are based on nonlinear partial differential equations (PDEs). This approach has an advantage and disadvantage compared to particle-based simulations, such as molecular dynamics and Monte Carlo simulations that are widely used to describe crystalline structures. The advantage is that the system is expressed using a continuum (mostly scalar) field in which the position of a *particle* is identified by a peak of the field. Because the density field is defined everywhere in the system, it is easier to analyse structures (symmetry, length scale, and so on) and deviation from a perfect crystalline structure (dislocation and disclination). These topological defects have been studied not only in solids[113], but also in soft materials[114, 115]. The dynamics of defects has been studied by using PDEs[108, 112]. The PFC is a natural extension of the conventional phase-field model, and is capable of describing diffusive time scale as well as local crystalline order[111]. Therefore, the PFC captures both elasticity and plasticity, and long-time behaviours of defect dynamics. The disadvantage is that there is no explicit form of interaction potential between particles, and efforts are necessary to map molecular interactions to the associated linear and nonlinear terms in PDEs[116].

The extension of the PFC (or SH) has been proposed in several ways. One is to introduce two length scales to destabilise the uniform state[117, 118]. This results in stable quasi-crystalline patterns such as decagonal (10-fold) and dodecagonal (12-fold) quasi-crystals in two dimensions. This model also reproduces an icosahedral pattern in three dimensions[119] although in this case, the two length scales are not linearly unstable but close to neutral stability. The quasi-crystal pattern appears due to a nonlinear selection mechanism. The second type of extension is to use many order parameters as a model of a multi-component system. Along this line, quasi-crystal patterns are reproduced in two dimensions[120] and in three dimensions[121]. The third type of extension is the weak crystallisation theory based on the Landau-Brazovskii theory [122]. The simplest version of this theory is nothing but the SH equation. In the weak crystallisation theory, anisotropic nonlinear terms are included to make complex patterns to reproduce, for example, quasi-crystals in two dimensions[123]. The model using the anisotropic nonlinear terms corresponds to an anisotropic interaction in a microscopic model, for example, using Janus particles, patchy particles[124], or polyhedral-shape particles[125].

### B. Inverse structural optimisation

Studies on inverse structural design have been performed mostly using particle-based models. The two major directions of the research is a control of the structure through the external template, and the design of potential interaction between particles. The template method is more intuitive than the potential design, but it requires a good template beforehand. On the other hand, the design of potential interactions does not need the prescribed structure, and the structure appears spontaneously from a uniform disordered state. This self-assembly of discrete particles has an advantage against the template approach because after optimisation, the structures are reproduced without external aids, and thus it is easy to scale up the structures[126–128]. Most of the studies along this line are discrete particles using either Monte Carlo, Molecular Dynamics simulations, or an extension of them. Examples include

simulated annealing for the parameters in Lennard-Jones-type isotropic interactions using the difference of the particle position from the target lattice in real space as a cost function [129, 130] (See also [131, 132] and references therein). These studies initiate the inverse structural design of the self-assembly of materials [133, 134]. The design is based on various optimisation techniques such as gradient method [135], generic algorithm [136], swarm optimisation [137, 138], the Covariance matrix adaption-evolutionary strategy (CMA-ES) [139, 140] and statistical physics-inspired inverse design (SP-ID) [140, 141].

Recently, the optimisation of interaction potentials by minimisation of the relative entropy combined with Monte Carlo simulations has been proposed [135, 142–146]. In those studies, parameters in a potential function are optimised by a gradient of pair-potential with respect to the parameters using the difference between the target and target radial distribution functions. The target structure is given by positions of particles, and by adding fluctuation around the positions, the radial distribution function for the structure is obtained. This method treats the probability of the microscopic structures for given parameters by the Boltzmann distribution. The inverse structural design has been applied to various soft materials such as colloidal assembly [135, 142–144, 146], block copolymers [137, 138], and granular media [147]. These methods, thus far, are the point estimate in which only the most likely parameters are estimated. In the BM-PDE, we compute the posterior distribution from which the uncertainty of the parameters can be evaluated. We also optimise the noise amplitude described by the inverse temperature  $\beta$ . This optimisation has not been performed in the previous works, and it is essential to estimate the best parameters for the target pattern with noise and without ground truth parameters. Lack of optimisation of the noise amplitude results in overfitting of the parameters for the target under noise [148]. Besides, the BM-PDE is an inverse structural design of the model described by PDEs. The optimisation problem of PDEs has focused on time-series data, and therefore, for the snapshot of a stationary structure, the inverse problem has not been formulated.

### C. parameter estimation of governing equations

The automatic discovery of the governing equations from data is a relevant topic in data-driven science. Parameter estimation of nonlinear PDEs has been studied [149–151] along the line of system identification, optimisation, and control of nonlinear dynamical systems [152]. Initiated by the success of physical laws for double-pendulum [153, 154], it is becoming feasible to estimate an equation of motion from data  $\psi$ . Recently, the method of sparse linear regression was successfully applied to ordinal differential equations (ODEs) [155] and partial differential equations (PDEs) for time-series data [156, 157]. The key idea is to minimise the error of  $\|\partial_t \psi(t) - f(\psi(t))\|$  under a certain norm  $\|\cdot\|$  with a regularisation term. The function  $f(\psi)$  is expanded in terms of polynomials of  $\psi$  with their coefficients. The polynomials may be replaced by a list of candidate terms in  $\psi$ . Then, the problem reduces to an estimation of the coefficients. Estimation based on sparse regression helps to make many coefficients zero, so that only a few terms remain in the estimated governing equation. Parsimony is an underlying philosophy of the method; the governing equation *should* be described by minimal terms in  $\psi$  and minimal terms in spatial derivatives in the case of PDEs.

Our problem to estimate the governing PDE from a given target pattern is fundamentally different from those studies in several respects. First, the methods mentioned above are based on the regression for the time-series data and therefore demands an accurate observation of  $\partial_t \psi$  [150, 152, 158, 159]. The target pattern in our problem is only one snapshot ideally satisfying  $\partial_t \psi = 0$  without noise, and therefore, we cannot use optimisation with respect to  $\partial_t \psi(t) - f(\psi(t))$ . Moreover, the information on the snapshot is far less than time-series data because, in the latter case, there are data of  $\psi(t)$  under different time  $t$ . Second, our interest is to estimate not only parameters but also the best model. Third, our target pattern is not necessarily produced from a numerical result but synthesised by a function of equation (14) that is independent of the models.

Another feature of BM-PDE is uncertainty quantification. The statistical inference has been widely used to estimate parameters with their errors (uncertainty) [160, 161]. This approach is well established in the linear regression problems. However, fewer studies have been made in uncertainty quantification for the estimation of governing equations, particularly for nonlinear PDEs [162–165]. The majority of the methods are based on point estimation in which parameters are estimated by minimising the cost function (energy) [155–157]. To estimate parameters in ODEs, statistical inference has been used for time-series data using approximate Bayesian computation [166].

### D. State-space model

To summarise the related studies on the parameter estimation of the dynamical equations, it is instructive to consider the state-space model. We first explain the estimation from time-series data, and then discuss the difficulty

in estimating the stationary problem. In the state-space model, the cost function is given by

$$E = \frac{1}{2} \sum_{i,\alpha} (\psi^*(\mathbf{x}_i, t_\alpha) - \psi(\mathbf{x}_i, t_\alpha))^2 + \frac{\lambda}{2} \int (\partial_t \psi(\mathbf{x}, t) - f_\mu(\psi(\mathbf{x}, t)))^2 d\mathbf{x}dt, \quad (\text{S1})$$

where the indices  $i$  and  $\alpha$  are measurement points in space and time, respectively (see also equation (1) in the main text). Within this model, the data at hand is  $\psi^*$ , and the model is parametrised by  $\mu$ . The first term in equation (S1) makes the pattern  $\psi(\mathbf{x}, t)$  closer to the data (measurement). The second term penalises the pattern that deviated from the model  $f_\mu(\cdot)$  under the parameters  $\mu$  (model). Here,  $\lambda$  sets a balance between the two terms in equation (S1). The cost function equation (S1) implies that both measurement and model have noise,

$$\psi^*(\mathbf{x}, t) = \psi(\mathbf{x}, t) + \xi_1 \quad (\text{S2})$$

$$\partial_t \psi(\mathbf{x}, t) = f_\mu(\psi(\mathbf{x}, t)) + \xi_2. \quad (\text{S3})$$

When the noise  $\xi_1$  and  $\xi_2$  are taken from the normal distribution, the likelihood is given by  $e^{-E}$ .

When the model is expressed by a deterministic equation, the trajectory of the solution  $\psi(\mathbf{x}, t)$  is described by the flow map from the initial condition,  $\psi(\mathbf{x}, t) = \Phi_t(\psi_0(\mathbf{x}))$ . In this case, the cost function is  $E[\psi^*(\mathbf{x})|\psi_0(\mathbf{x}), \mu]$ [167]. Then, using the Bayes' theorem, we may consider the posterior distribution from three choices. The first choice is the posterior distribution of parameters  $p(\mu|\psi^*(\mathbf{x}), \psi_0(\mathbf{x}))$  from the prior distribution of the parameters  $p(\mu)$  with a fixed initial condition. The second choice is  $p(\psi_0(\mathbf{x})|\psi^*(\mathbf{x}), \mu)$  from the prior distribution of the initial conditions  $p(\psi_0(\mathbf{x}))$  with a fixed parameter. The third choice is  $p(\psi_0(\mathbf{x}), \mu|\psi^*(\mathbf{x}))$  from both the prior distribution of parameters and the initial conditions. Even in the first and the second choices, the parameters and the initial conditions may be estimated from hyper-parameter estimation[160]. In data assimilation, the initial condition is often estimated, and the parameters are estimated by hyper-parameter estimation (second choice)[168]. When both the initial condition and the parameters are estimated (third choice), the state augmentation has been used [169]. In this method, the parameters are treated as state variables. In [170], parameters in the phase-field model are estimated, and the initial condition is estimated by hyper-parameter estimation.

All the choices discussed above estimate the unique initial condition. However, for the stationary data in pattern formation, the same pattern appears from different initial conditions. It is rather desired that the target pattern can be generated from a wide range of the initial conditions. In our method, we consider the posterior distribution of parameters  $p(\mu|\psi^*(\mathbf{x}), \psi_0(\mathbf{x}))$  with a fixed initial condition (first choice), but the posterior distribution is marginalised about the initial conditions. This marginalisation is a core of our approach, and it enables us to avoid non-unique estimation for stationary data, and to estimate the model that generates the target pattern as a stable state.

When there is no measurement noise, that is  $\xi_1 = 0$ , and the measurement is dense in time so that the time derivative  $\partial_t \psi$  is accurately measured (see Fig. S1(a)), we may drop the first term in equation (S1), and the cost function becomes

$$E_{\text{reg}} = \frac{1}{2} \sum_{i,\alpha} (\partial_t \psi(\mathbf{x}_i, t_\alpha) - f_\mu(\psi(\mathbf{x}_i, t_\alpha)))^2. \quad (\text{S4})$$

We call this cost function the regression model, which has been used to estimate  $\mu$  for a PDE model in [149, 150, 152]. By adding the  $L^1$  regularisation term in  $E_{\text{reg}}$ , the sparse estimation of the parameters was demonstrated in [155–157, 171]. With this regularisation, only a few parameters become nonzero. This sparse regression method is appealing because even when the dynamical equation  $f(\psi)$  is nonlinear, the model can be parameterised by linear terms in the parameters such as  $\mu_2 \psi^2$ ,  $\mu_{2,1} \nabla \psi^2$ , and so on[157]. Therefore, the estimation of the parameters falls into the linear regression problem, which can be handled in various ways[172]. The uncertainty of the estimation is also obtained once the linear regression problem is reformulated in the Bayesian framework[161, 173]. Sparsity may be included in the prior distribution[171]. Nevertheless, in the regression model, the error is included only in the model. Therefore, when there is measurement noise in data, the estimation does not work even for the linear regression problem[152]. This issue has been studied by the total least square[152].

When the measurement is sparse in time,  $\psi$  in equation (S1) is available only at sparse points as  $\psi^*$  (Fig. S1(b)). In this case, both the parameters  $\mu$  and the unknown solution of the PDE  $\psi(\mathbf{x}, t)$  need to be estimated. The simplest way is to interpolate the unobserved data in time and space, for example, by the spline method[174, 175]. Clearly, to interpolate accurately, the observation point should be dense if not complete, and noise should be small[149, 150]. If the interpolation works well, the estimation of the parameters can be made in the same way as the above-mentioned regression method.

When the observation is so sparse that the interpolation does not work, both the parameters  $\mu$  and the solution of the PDE  $\psi(\mathbf{x}, t)$  must be estimated. Even if the observation is linear as  $\psi^* = \psi + \xi_2$ , and the model is assumed to be

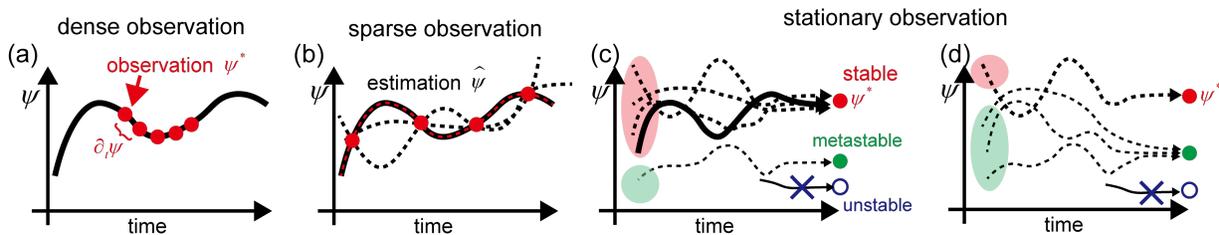


FIG. S1. Schematic estimated trajectories for dense observation (a), sparse observation (b), and stationary data (c,d). The observations  $\psi^*$  are shown in red points. The true trajectory is shown in the solid curve, whereas the estimated trajectories  $\hat{\psi}$  are shown in dashed curves. For the stationary data, two models are shown. (c) The stationary data is generated from broad initial conditions. (d) The stationary data is generated from narrower initial conditions. In this case, most of the initial conditions lead to the meta-stable state. In both cases, the unstable solution cannot be achieved from the initial conditions, even though it satisfies  $f(\psi) = 0$ .

linear in the parameters, the estimation for the sparse observation is inevitably nonlinear. This is in contrast with the regression method discussed above. The reason is that the dynamical equation  $f_\mu(\psi)$  is nonlinear, and the solution of the PDE is a nonlinear function of the parameters  $\mu$ . Therefore, the estimation under the sparse observation is far more difficult than that under the dense observation. To handle this issue, various methods have been proposed. Below, we discuss these methods briefly, and point out the difficulty to apply them for stationary data.

To our knowledge, all the studies on the estimation of governing equations described by PDEs are based on time-series data. Our problem is to estimate the best model and parameters from one snapshot of a pattern at the stationary state. Trajectory matching cannot be used for this problem because there is no information about the initial condition. We need to eliminate the dependence on the initial condition. To do this, we marginalise the initial condition, which is taken from a random distribution. We also have to identify the same patterns which have a different orientation. The different orientation arises from different initial conditions under the same model and the same parameters. This argument is the reason why the order parameter has to be introduced in BM-PDE.

### 1. Gaussian process

First of all, when both the observation and the model are linear, the Gaussian process is a very powerful tool[176]. Because the time and spatial derivatives are linear operations, the pattern  $\psi(\mathbf{x}, t)$  taken from the Gaussian prior distribution remains Gaussian. Therefore, the posterior distribution may be computed by inversion of the matrix of the kernel. This method has been used for linear ODEs and PDEs [177–180]. The method using the Gaussian process is robust against noise[177, 180], and can handle hidden variables[178], and boundary conditions[179], but the method is limited to the linear models. For nonlinear problems, the Gaussian process may be used by marginalising the time derivative of the variables[181–183]. When the time derivative of the solution  $\psi(\mathbf{x}, t)$  can be measured in the dense observation, the model can be linearised between the small time steps between the two observations. In this case, the Gaussian process works even for a nonlinear model[184]. The drawback of the Gaussian process is the interpretability of the estimated model. The estimated model can reproduce the observation, but its physical meaning is hardly understood from the estimation results. Another issue is a computational cost. For the PDEs, the method assigns the Gaussian prior at each spatial and time point. The matrix inversion requires  $\mathcal{O}(M^3)$  computation where  $M$  is the total spatial points  $M \sim N^d$  where  $N$  is the number of mesh points in each spatial direction. Therefore, for higher dimensional systems, this method is not feasible.

### 2. shooting method

The second method is the shooting method (also multiple shooting and extended multiple shooting method)[150, 185, 186]. The idea is to solve equation (S1) explicitly by using the measured data as an initial condition. In most cases, solving the model is performed numerically. The estimated solution of  $\psi(\mathbf{x}, t)$  is obtained under the given parameters. By minimising the first term in equation (S1) with respect to the parameters, the best parameters may be estimated.

In the simplest shooting method, we use only one measurement for the initial condition and use the rest of the data to compute the cost function. When the model is sensitive to the initial condition, such as chaotic dynamical systems,

the error grows exponentially in time. Therefore, this method results in a larger error away from the initial condition. To handle such data, in the multiple shooting method, several data are used as initial conditions, and estimate  $\psi$  in a short period of time for each of the initial conditions.

### 3. filtering and sequential data assimilation

Sequential data assimilation is widely used to fit data to a model at hand [167, 169]. It consists of three steps: prediction, filtering, and smoothing. In prediction, the model is evaluated, and the state at the next step in time  $\psi(\mathbf{x}, t + \Delta t)$  is predicted from the current time  $\psi(\mathbf{x}, t)$ . In filtering, the predicted value  $\psi(\mathbf{x}, t + \Delta t)$  is revised to fit the measurement  $\psi^*$ . These steps are carried out sequentially in time. Finally, in smoothing, the estimated  $\hat{\psi}(\mathbf{x}, t)$  is further revised to fit all the past measurement.

For the deterministic model, the value of  $\psi(\mathbf{x}, t)$  at any time  $t$  may be expressed by the flow map  $\psi(\mathbf{x}, t) = \Phi_t(\psi_0)$  for the initial condition  $\psi_0$ . Therefore, sequential assimilation is often used to estimate the initial conditions. The estimation of the parameters can also be included in the sequential data assimilation.

### 4. adjoint method and sensitivity analysis

The adjoint method is called the variational method or 4DVAR in data assimilation. The main idea is to introduce a Lagrange multiplier  $\lambda$ , and define the cost function as

$$E_{adj} = \frac{1}{2} \sum_{i,\alpha} (\psi^*(\mathbf{x}_i, t_\alpha) - \psi(\mathbf{x}_i, t_\alpha))^2 + \int \lambda(\mathbf{x}, t) (\partial_t \psi(\mathbf{x}, t) - f_\mu(\psi(\mathbf{x}, t))) d\mathbf{x}dt. \quad (S5)$$

The variation of the cost function with respect to the parameters results in the gradient dynamics of both the parameters and the Lagrange multiplier. The equation of the Lagrange multiplier is nothing but solving the model backwards to correct the error between the estimated trajectory and the measurement. This method was successfully applied to pattern forming PDEs [151, 165]. However, in this method, the initial condition has to be fixed because the method makes an estimation by solving the model in forward and backward directions in time. Estimating the initial condition may be possible by the hyperparameter estimation. We are not aware of the discussion about the marginalisation of initial conditions and the stability of the estimated model within the adjoint method.

## E. estimation of the model with a stationary state

When the observation is made only at the stationary state, the data does not have information about trajectories before reaching the stationary state (Fig. S1(c,d)). As we have seen in the previous section, the various methods for the parameter estimation of the state-space model rely on time series. Clearly, the method using the time derivative or its interpolation cannot be applied for the stationary problem. Only the stationary data is the *sparsest* measurement in time series. The shooting method, the sequential data assimilation, and the adjoint method essentially estimate trajectories between two measurements in time. In the stationary problem, there are many initial conditions leading to the same stationary state. We are interested in the model that can reproduce the stationary target pattern from a wider range of the initial conditions. Then, the estimated model is expected to reproduce the target pattern as a stable state.

Besides such obvious difference from the estimation from time series, the estimation from the stationary data has several difficulties. There are two types of problems for the estimation from the stationary data; the first is to find parameters  $\mu$  for

$$f_\mu(\psi^*) = 0. \quad (S6)$$

The second is to find the parameters for  $\psi^*$  as a solution of

$$\partial_t \psi = f_\mu(\psi) \quad (S7)$$

at  $t \rightarrow \infty$ . Our method, BM-PDE, focuses on the second problem. The advantage is that it is possible to make sure that the estimated model can reproduce approximately the target pattern as a stable solution. This is done by marginalising the initial conditions. When the solution is obtained from a wide range of the initial conditions, the model should reproduce the pattern in a stable manner.

The first problem does not care about the initial conditions. When the model  $f_\mu(\psi)$  is nonlinear, there may be several stable and unstable solutions. For this problem, it is not guaranteed whether the target pattern is a stable or unstable solution of the estimated model. In addition, the estimation of parameters is not unique; when  $f_\mu(\psi^*) = 0$ , a close neighbourhood,  $\mu + \delta\mu$ , of  $\mu$  may also satisfy  $f_\mu(\psi^*) = 0$  as long as  $f$  is not at a bifurcation point. To make the estimation unique, we need a regularisation term. The ridge regression ( $L^2$  regularisation) or sparse regression ( $L^1$  regularisation) may fix the estimation, but they are independent of the stability of the target pattern. Therefore, these regularisation methods do not guarantee that the estimated model reproduce the target pattern as a stable state.

### SIII. FAMILY OF MODELS

#### A. Phase-field crystal model and its generalisation

For a single wavelength, we have  $m = 1$ , and we recover the conserved SH equation (the simplest PFC model). In order to control the position and amplitude of unstable mode, we use the following form,

$$\mathcal{L}_k = -\frac{a_0}{q_0^4} k^2 (k^2 - 2q_0^2) - s_0 k^2 (q_0^2 - k^2)^2 \quad (\text{S8})$$

With this form, patterns with  $k = q_0$  appear, and its stability is denoted by  $a_0$  (see Fig. S2). Here,  $s_0$  expresses sharpness of the peak in the spectrum. For larger  $s_0$ , the peak becomes sharper (Fig. S2).

For two length scales, we use another length scale  $k = q_1$  in addition to  $k = q_0$ , and the linear stability around the uniform state for the two length scales,  $a_0$  for  $k = q_0$  and  $a_1$  for  $k = q_1$  as shown in Fig. S2. This model has been used to reproduce quasi-crystal patterns [119]. The explicit form of the linear operator is given by

$$\mathcal{L}_k = a_0 S_0(k) + a_1 S_1(k) - \frac{s_0 k^2}{q_1^4} (q_0^2 - k^2)^2 (q_1^2 - k^2)^2 \quad (\text{S9})$$

$$S_0(k) = \frac{1}{q_0^4 (q_0^2 - q_1^2)^3} k^2 (q_1^2 - k^2)^2 ((q_1^2 - 3q_0^2)k^2 + q_0^2(-2q_1^2 + 4q_0^2)) \quad (\text{S10})$$

$$S_1(k) = \frac{1}{q_1^4 (q_0^2 - q_1^2)^3} k^2 (q_0^2 - k^2)^2 ((3q_1^2 - q_0^2)k^2 + q_1^2(2q_0^2 - 4q_1^2)). \quad (\text{S11})$$

To make sharp enough peaks we use  $s_0 = 100$ .

For three length scales, we may extend the above-discussed expressions and obtain

$$\mathcal{L}_k = a_0 S_0(k) + a_1 S_1(k) + a_2 S_2(k) - s_0 k^2 (q_0^2 - k^2)^2 (q_1^2 - k^2)^2 (q_2^2 - k^2)^2 \quad (\text{S12})$$

$$S_0(k) = \frac{1}{q_0^4 (q_0^2 - q_1^2)^3 (q_2^2 - q_0^2)^3} k^2 (k^2 - q_1^2)^2 (k^2 - q_2^2)^2 \times (q_0^2(-6q_0^4 + 4q_0^2 q_2^2 + 4q_0^2 q_1^2 - 2q_1^2 q_2^2) + k^2(5q_0^4 - 3q_0^2 q_2^2 - 3q_0^2 q_1^2 + q_1^2 q_2^2)) \quad (\text{S13})$$

$$S_1(k) = \frac{1}{q_1^4 (q_0^2 - q_1^2)^3 (q_1^2 - q_2^2)^3} k^2 (k^2 - q_0^2)^2 (k^2 - q_2^2)^2 \times (q_1^2(-2q_0^2 q_2^2 + 4q_0^2 q_1^2 + 4q_1^2 q_2^2 - 6q_1^4) + k^2(q_0^2 q_2^2 - 3q_0^2 q_1^2 - 3q_1^2 q_2^2 + 5q_1^4)) \quad (\text{S14})$$

$$S_2(k) = \frac{1}{q_2^4 (q_2^2 - q_0^2)^3 (q_1^2 - q_2^2)^3} k^2 (k^2 - q_0^2)^2 (k^2 - q_1^2)^2 \times (q_2^2(-2q_0^2 q_1^2 + 4q_0^2 q_2^2 + 4q_1^2 q_2^2 - 6q_2^4) + k^2(q_0^2 q_1^2 - 3q_0^2 q_2^2 - 3q_1^2 q_2^2 + 5q_2^4)). \quad (\text{S15})$$

The linear spectrum is shown in Fig. S2.

#### B. Polynomial expansion of the linear operation with respect to wavenumbers

Our family of models is based on the number of length scales. This is demonstrated in the functional form of the linear operator in the Fourier space shown in Fig. S2 and equations (10)-(11). The benefit of this approach is that parameters  $\mu$  have clear physical meaning: the wavenumbers  $q_i$  correspond to the characteristic length scales and stability at the wavenumbers  $a_i$ . The disadvantage of this approach is that we have to treat the different number of

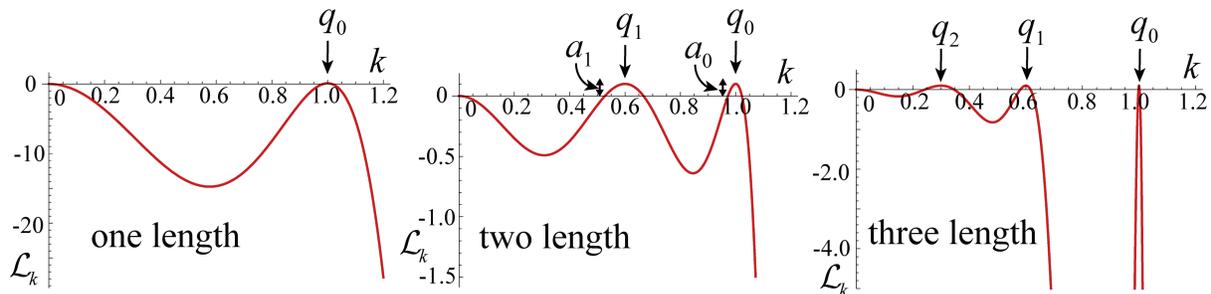


FIG. S2. Linear stability of multi-length-scale phase-field crystal equations. The plots show eigenvalues as a function of the wavenumber. Positive eigenvalues imply the uniform state is unstable.

length scales independently. This implies that the two-length-scale model  $m = m_2$  does not include the one-length-scale model. If  $q_0$  and  $q_1$  are overlapped in  $m = m_2$ , the operator  $\mathcal{L}$  diverges, as seen in equation (S9). On the other hand, we may consider a family of the model by expanding  $\mathcal{L}_k$  in equation (11) by a polynomial expansion of  $k$ . In this section, we discuss the drawbacks of this approach and explain why we use our family of models by equation (11).

Within the approach of polynomial expansion, parameters  $\{\mu\}$  are chosen as coefficients of a polynomial expansion of  $\mathcal{L}_k = \sum_i \mu_i k^i$ . The advantage of this expansion is that we may express  $\mathcal{L}_k$  of different models with the different number of length scales in Fig. S2 in a unified way. Once we truncate the expansion up to  $k^{14}$ , we may express a three-length-scale model, and when the coefficients of  $k^{12}$  and  $k^{14}$  are identically zero, the model describes two length scales. We may use sparse regression by adding regularisation and use a particular type of prior distribution of the parameters[187].

The disadvantage of this method is that the coefficients of the expansion in terms of polynomials of  $k$  do not have explicit physical meanings. In practical term, the range of the parameters is too broad, so that estimation is not possible. For example, the numerically produced dodecagonal quasi-crystal discussed in Sec. SIV A is expressed as

$$\mathcal{L}_k \approx -99.7k^2 + 945.6k^4 - 2984.8k^6 + 3531.8k^8 - 1392.8k^{10}; \quad (\text{S16})$$

and the precision of 0.1 is necessary to obtain quasi-crystal. Therefore, the range of the parameters should be  $[-4000.0, 4000.0]$ . Practically, it is impossible to get a reasonable estimate from a such wide range of the parameters, and even if possible, uncertainty is certainly much larger than 0.1, and thus there is no guarantee to obtain quasi crystal within the error.

Finally, we should point out that the sparse regression does not work when parameters are correlated [188]. In the system that we are studying, the parameters must be correlated because the amplitude and position of the unstable wavenumber are set by the peak of  $\mathcal{L}_k$ , and it is determined by the balance of several terms in the polynomial expansion.

### C. Expansion of the linear operator by orthogonal bases

In [151, 165], it was proposed that the directed correlation function is expanded by the Hermite polynomials. This method is appealing because the polynomial forms orthogonal bases, and therefore, its coefficient must be independent. In our formula, this method corresponds to different parametrisation of the linear operator  $\mathcal{L}_\mu^{(m)}$  in equation (5), or  $\mathcal{L}_k$  in equation (10), by the Hermite polynomials as equation (S17). Instead of (11), we may expand the linear operator in (6) and (10) as

$$\mathcal{L}(k) = -a_0 k^2 + k^2 \sum_{n=1}^{n_{\max}} a_n \frac{e^{-\frac{1}{2} \frac{k}{k_0}}}{(2^n n! \sqrt{\pi})^{1/2}} H_n(k/k_0) \quad (\text{S17})$$

where  $H_n(x)$  is the Hermite polynomial. The reference wavenumber is denoted by  $k_0$ . We set to be  $k_0 = 0.25$ , and  $n_{\max} = 10$ . We note that the first term with  $a_0$  appears from the conserved system. The linear operator in the Fourier space diverges as  $\mathcal{L}(k \rightarrow \infty) \sim -k^2$ , which make sure that fluctuations at the higher wavenumbers decay quickly. In the method shown in the previous section, the scale of the divergence is different, for example,  $\mathcal{L}(k \rightarrow \infty) \sim -k^{10}$ , but the qualitative shape of the linear spectrum of the two methods are the same (see Fig.3(g) and Fig.S3(e)).

For the target pattern shown in Fig.S3(a,b), we estimate the parameters in equation (S17). The estimated parameters are shown in Fig.S3(f). Most of the parameters are close to zero, and there are two finite modes. Using the estimate parameters, we may generate quasi-crystal as shown in Fig.S3(c). The estimated linear spectrum is shown in Fig.S3(e). The qualitative shape of the linear spectrum is the same as Fig.3(g) used in the previous section. In fact, there are two peaks whose ratio is  $1/(2 \cos(\pi/12))$ . The ratio is consistent with the result in the previous sections.

The benefit of the expansion by the basis functions is that there is no need to consider different models with the different number of length scales. The arbitrary number of length scales can be expressed by the appropriate superposition of the bases. The disadvantage of this method is that the position of the characteristic wavelength, characterised by the peak in the linear spectrum, cannot be imposed as a parameter.

#### D. design of nonlinear terms

Our approach is not limited to the estimation of linear parts in the model. Instead of equation (5), we may consider the two-component model described by the two density fields  $\psi_1(\mathbf{x})$  and  $\psi_2(\mathbf{x})$ . Both the fields have their own single length scales  $q_1$  for  $\psi_1(\mathbf{x})$  and  $q_2$  for  $\psi_2(\mathbf{x})$ , which may be different from each other. The model is expressed by

$$\partial_t \psi_1(\mathbf{x}) = \mathcal{L}_{\mu_1} \psi_1(\mathbf{x}) + \Delta \mathcal{N}_1[\psi_1, \psi_2] \quad (\text{S18})$$

$$\partial_t \psi_2(\mathbf{x}) = \mathcal{L}_{\mu_2} \psi_2(\mathbf{x}) + \Delta \mathcal{N}_2[\psi_1, \psi_2], \quad (\text{S19})$$

with the nonlinear terms  $\mathcal{N}_1[\psi_1, \psi_2]$  and  $\mathcal{N}_2[\psi_1, \psi_2]$  are given by polynomials of  $\psi_1$  and  $\psi_2$ , and are parametrised by  $b_n$  as

$$\mathcal{N}_1[\psi_1, \psi_2] = \psi_1^3 - 2b_1\psi_1\psi_2 - b_2\psi_2^2 - b_3\psi_2^3 - 2b_4\psi_1\psi_2^2 - 3b_5\psi_1^2\psi_2 \quad (\text{S20})$$

$$\mathcal{N}_2[\psi_1, \psi_2] = \psi_2^3 - b_1\psi_1^2 - 2b_2\psi_1\psi_2 - 3b_3\psi_1\psi_2^2 - 2b_4\psi_1^2\psi_2 - b_5\psi_1^3. \quad (\text{S21})$$

The first terms in (S20) and (S21) are the same as the families of models in the previous sections. For simplicity, we assume the gradient structure of the model, that is, there exists the free energy functional  $F$  so that the model is expressed by  $\partial_t \psi_{1,2} = \Delta \delta F / \delta \psi_{1,2}$ . The free energy functional contains the terms of coupling between  $\psi_1$  and  $\psi_2$  up to the fourth-order of the polynomials. In this model, the parameters are the system size in  $x$  and  $y$  directions, mean density  $\bar{\psi}$  which is assumed to be the same for the two fields, the two wavenumbers  $q_1$  and  $q_2$ , and corresponding linear spectrum  $a_1$  and  $a_2$ , and the coefficients  $b_1, \dots, b_5$  of nonlinear coupling terms. Here,  $q_1$  is fixed to be 1 without loss of generality. We use the periodic boundary conditions.

The goal of the estimation in this model is to specify the most relevant nonlinear term among the candidates. To do this, we use sparse Bayesian formula[189]. Instead of the uniform prior distribution, we use the prior distribution  $\exp(-\kappa \sum_{i=1}^5 |b_i|)$  for the coefficients of the nonlinear terms. Due to the prior distribution, only a few of  $b_i$  becomes nonzero. We have performed the parameter estimation both with and without sparsity.

We use the same target pattern shown in Fig. S3(a,b) as used in Fig. 4(c), and then perform the parameter estimation. The generated pattern from the estimated parameters is shown in Fig. S3(d). The pattern has dodecagonal symmetry as in the target pattern. The posterior distributions of the parameters are shown in Fig. S3(h,i). Without sparsity, all the nonlinear coupling terms contribute to the estimation. On the other hand, with sparsity, only the term  $\psi_2^3$  in the  $\mathcal{N}_1[\psi_1, \psi_2]$  is nonzero, and other coefficients become close to zero. Therefore, the term  $\psi_2^3$  is the most relevant nonlinear term.

Surprisingly, the estimated wavenumber is  $q_2 \approx 1/\sqrt{2}$  with respect to  $q_1 = 1$ . This ratio  $q_2/q_1$  is different from that in the two-length-scale model used in the previous section ( $q_0/q_1 \approx 1/(2 \cos(\pi/12))$ ). This can be understood by a quadrilateral in the Fourier space shown in Fig. S3(g). In the linear spectrum, the modes at the wavenumbers  $q_1$  and  $q_2$  dominate. The nonlinear terms affect the pattern selection when triangle (second-order terms of  $\psi_{1,2}$  in equations (S20) and (S21)) or quadrilateral (third-order terms of  $\psi_{1,2}$ ) appear to connect the spots in Fig. S3(g). In the previous section, the nonlinear term  $\bar{\psi}\psi^2$  determines the stability of the dodecagonal quasi-crystal. To draw a triangle to connect the spots in the inner and outer circles in Fig. S3(g), the angle of the triangle should be  $\pi/6$ . This is the reason why the ratio  $q_0/q_1 = 1/(2 \cos(\pi/12))$  appears [117].

In the two-component system given by equations (S20) and (S21), the dominant nonlinear term is third-order  $\psi_2^3$  in  $\mathcal{N}_1[\psi_1, \psi_2]$  ( $\psi_1\psi_2^2$  in  $\mathcal{N}_2[\psi_1, \psi_2]$ ). Therefore, a quadrilateral consists of four wave vectors  $\mathbf{q}_1, \dots, \mathbf{q}_4$  make a contribution when  $\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4 = 0$ . From the geometry shown in Fig. S3(g), the ratio between the two wavenumbers must be  $1/\sqrt{2}$ . To our knowledge, this is the first dodecagonal quasi-crystal generated from PDEs with the ratio  $1/\sqrt{2}$ .

Figure. S3(j) shows the histogram of the cost function of the generated patterns from the models only with the nonzero coefficient of the coupling term  $b_1$  or  $b_5$  in equations (S20) and (S21). The parameters are fixed in each model, but each stationary pattern is generated from a different initial condition. Both models can generate the

DDQC patterns in two dimensions. The cost function of the DDQC is  $E \simeq 10$ . The histogram shows that the model with  $b_5 \neq 0$  is more stable in the sense that the DDQC is generated from a wide range of initial conditions. Even though the model with  $b_2 \neq 0$  can generate the DDQC, most of the initial conditions lead to the hexagonal patterns. The result demonstrates that our method may choose parameters so that the model can reproduce the target pattern from a wide range of initial conditions.

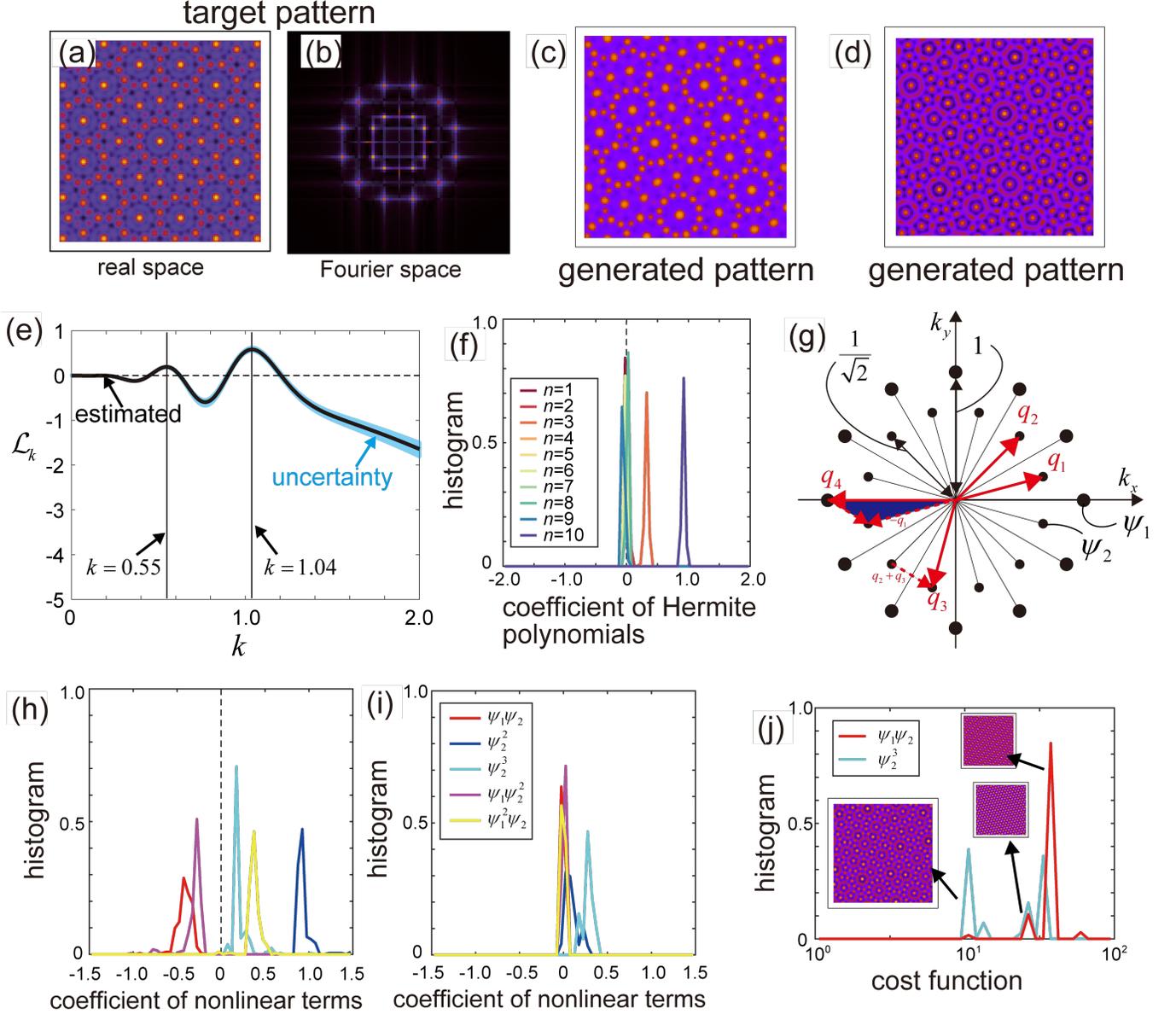


FIG. S3. (a,b) The target pattern of the synthesised dodecagonal quasi-crystal (a) and the intensity of its Fourier transformation (b). (c,e,f) The estimation for the model using Hermite polynomials. (c) The generated pattern for the model using the Hermite polynomials. (e) The linear spectrum (black lines) as a function of wavenumber for the estimated parameters for the Hermite polynomials. The uncertainty of the estimation is shown by the range in light blue. The two solid black lines show the wavenumbers at the peaks. (f) The histogram of the estimated coefficients of the Hermite polynomials. (d,g-j) The estimation for the two-component model. (d) The generated pattern from the estimated parameters for the two-component model. (g) The schematic illustration of the intensity in the Fourier space of the estimated dodecagonal quasi-crystal. The four vectors  $\mathbf{q}_1$ - $\mathbf{q}_4$  are the example of the dominant coupling term. The blue triangle corresponds to the three vectors  $\mathbf{q}_1$ ,  $\mathbf{q}_4$ , and  $\mathbf{q}_2 + \mathbf{q}_3$ . (h,i) The estimated parameters of the nonlinear coupling terms between the two-component  $\psi_1$  and  $\psi_2$  with (i) and without (h) sparsity. The colour legend indicates the nonlinear terms for  $\psi_1$ . (j) The histogram of the cost function under the model only with  $\psi_1\psi_2$  (black) and  $\psi_2^3$  (red) in equation (S20).

## SIV. EXAMPLES

### A. Target pattern produced by a numerical simulation: two-dimensional dodecagonal quasi-crystal

The process of parameter estimation is shown in Fig. S4(a). At the initial parameters, estimation is poor in all temperatures,  $\beta$  (corresponding to the observation noise). After several MC steps, the cost function decreases drastically at the lower temperature, and reach its steady state. At the steady states, sampled patterns are very closed to the target pattern. At the higher temperature, the parameter change in each MC step is too large to converge to a unique pattern; rather, it exhibits nearly random motion in the parameter space. In Fig. S4(b), we demonstrate successful sampling in the parameter space by showing the cost function, or we call it energy, covers a wide range of its values. The distribution of the cost function has two regions connected by the energy gap. The lower cost function corresponds to a quasi-crystal, which is qualitatively similar to the target pattern, while the higher cost function is dominated by hexagonal patterns. This result implies that there are multiple minima of the cost function. This means that the standard MC simulation does not converge to the ground truth unless an initial parameter set is chosen nicely,

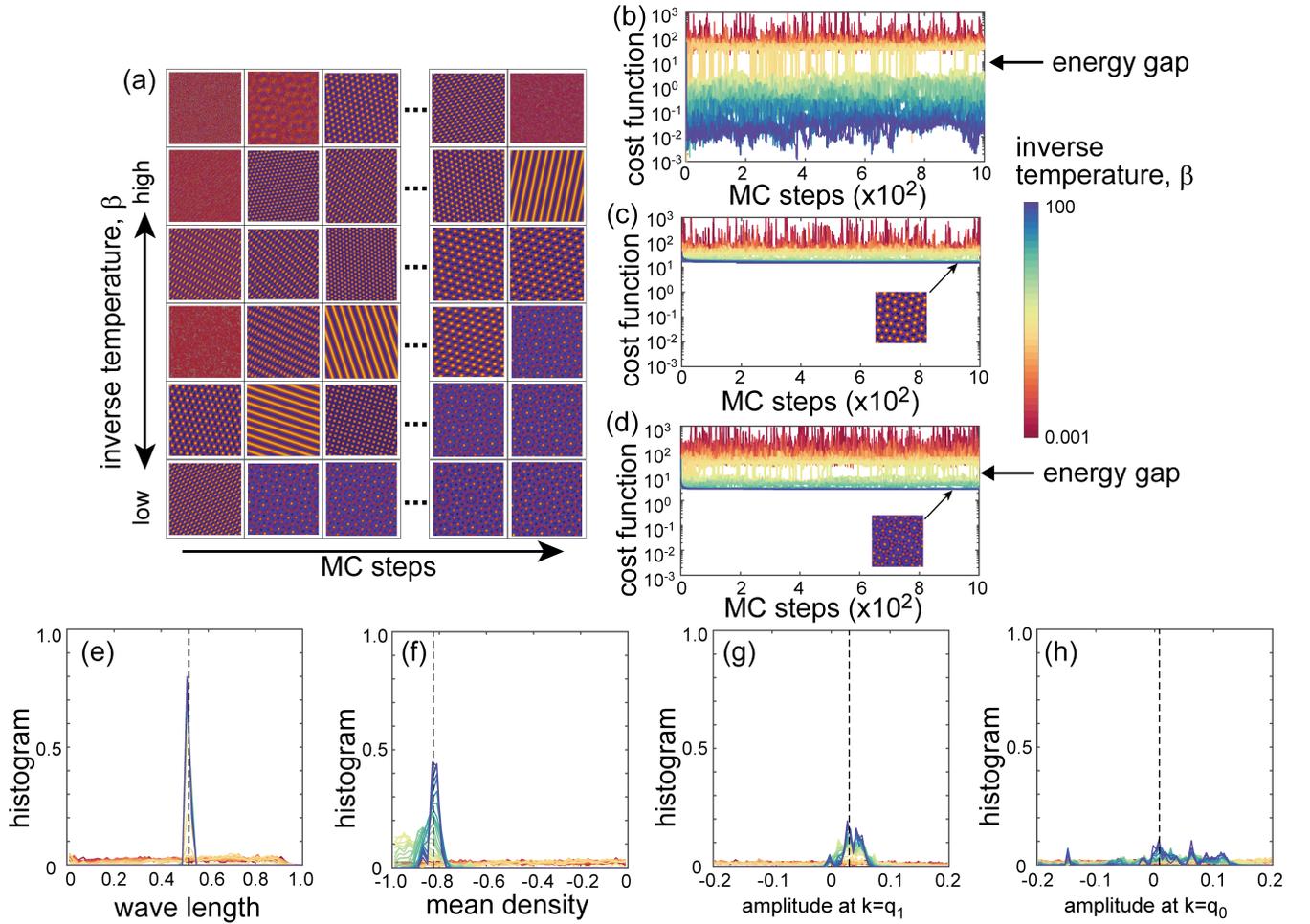


FIG. S4. (a) Snapshots of two-dimensional pattern with two length scales under different inverse temperature  $\beta$  during the Monte Carlo simulations for the target pattern of numerically produced dodecagonal quasi-crystal. (b-d) The cost function during the steps in replica-exchange Monte Carlo (REMC) simulations in the parameter space at the steady state with one (c), two (b), and three (d) length scales, respectively. (e-h) Histogram of the estimated parameters during REMC steps with vertical dashed lines indicating their ground truth. The different colours show different inverse temperature,  $\beta$ .

The same algorithm is used for the different models with one length scale or three length scales (Fig. S4(c,d)).

TABLE S1. The ground truth and estimated parameter values for numerically produced 12-fold quasi-crystal. Errors are evaluated from the standard deviation of each data.

parameters	ground truth	estimated (one length)	estimated (two length)	estimated (three length)
$dx$	1.020	$0.987 \pm 0.00615$	$1.0217 \pm 0.0024$	$1.003 \pm 0.00033$
$dy$	1.026	$0.994 \pm 0.0135$	$1.0232 \pm 0.00319$	$1.0040 \pm 0.0033$
$\bar{\psi}$	-0.826	$-0.389 \pm 0.0272$	$-0.820 \pm 0.0163$	$-0.682 \pm 0.018$
$a_0$	0.00794	$0.055 \pm 0.0113$	$0.0391 \pm 0.0594$	$-0.150 \pm 0.0081$
$q_1$	0.518	-	$0.518 \pm 0.00214$	$0.729 \pm 0.0023$
$a_1$	0.0306	-	$0.0375 \pm 0.0140$	$0.0383 \pm 0.020$
$q_2$	-	-	-	$0.368 \pm 0.0018$
$a_2$	-	-	-	$-0.136 \pm 0.011$

Both REMC simulations converge to their steady histograms of the cost function, both of which are higher than the histogram of the two-length-scale model. This result means the two-length-scale model is indeed most likely for the target pattern. In fact, the marginal likelihood, which we also call free energy  $F(\beta, m)$ , shown in Fig.3(e) in the main text has the lowest value for the two-length-scale model  $m = m_2$ . Because this target pattern has corresponding ground truth parameter values and the model of equation (5) is deterministic, the minimum of the free energy is, strictly speaking, achieved at  $\beta \rightarrow \infty$ . Nevertheless, we obtain a clear deviation of the free energy of the wrong models  $m = m_1$  and  $m = m_3$ , as they increase at large  $\beta$ . This is because the wrong models do not have exactly the same pattern as the target pattern in their solution, and therefore, the best estimation is made at the finite noise level, resulting in the minimum at the finite  $\beta$ . In those models, estimation of parameters is carried out at the optimal temperature  $\hat{\beta}$  at which the marginal likelihood has its minimum value.

The probability of each model is calculated by marginalising the probability  $P \propto e^{-F[\beta]}$  at each temperature,  $\beta$ , about the whole temperature range. This probability is also well captured by the minimal value of the marginal likelihood (free energy) of each model. The result is shown in the inset of Fig.3(e). In this example, the probability of selecting the two-length-scale model is almost 100%.

The estimate parameters are shown in Table S1, together with the uncertainty of the estimation. The uncertainty quantification is made by the standard deviation of the posterior probability distribution of the parameters around their mean value. The ground truth is indeed within the error.

Histograms of the parameters at each  $\beta$  under  $m = m_2$  are shown in Fig.S4(e-h). As we can see, at a lower temperature (higher  $\beta$  in the blue line), the distribution accumulates around the ground truth whereas, at the higher  $\beta$  (low  $\beta$  shown in the red line), the distribution covers the whole range of the parameters. From the high  $\beta$  to the low  $\beta$ , the distribution gradually becomes shrinking. This suggests that RMEC can well sample the true parameter distribution.

The width of the histograms at the low  $\beta$  varies from parameter to parameter. For the wavelength  $q_1$ , the distribution is narrow around the ground truth, as demonstrated in the small standard deviation in Table S1. The mean density  $\bar{\psi}$  and the spectrum amplitude at  $q = q_1$  (see Fig. S2) has broader distribution, but still, there is a clear peak near the ground truth. On the other hand, the distribution of the spectrum amplitude at  $q = q_0$  is very broad even at the lowest  $\beta$ . In fact, the error in the estimation of  $a_0$  is comparable to its estimated value, suggesting that we have a poor estimation on  $a_0$ . We suspect that this is because the model does not care about the value of  $a_0$  once the mode at the wavelength  $q_1$  becomes unstable as its amplitude  $a_1 > 0$ . Due to nonlinear interaction between different modes, even the mode with a negative spectrum, which is linearly stable, becomes unstable and grows. Therefore, the mode at the wavelength  $q_0 = 1$  may have arbitrary amplitude as long as its spectrum is close to zero that it can be destabilised by the mode at  $q_1$ .

The estimated parameters in the one-length-scale and three-length-scale models deviate from the ground truth. One length scale is not enough to reproduce a quasi-crystalline pattern, and therefore, the estimated parameters correspond to hexagonal patterns. The three-length-scale model, on the other hand, does reproduce a quasi-crystalline pattern, which is similar to the target pattern. Therefore, the estimated parameters, particularly the wavelength  $q_1$  and  $q_0$  satisfy the ratio  $q_0/q_1 \simeq 1.997$ , which is close to  $2 \cos(\pi/12)$ . Nevertheless, the cost function is higher than that of the two-length-scale model because the three-length-scale model is not a true model, and the sampled pattern is quantitatively different from the target pattern. We note that the cost function of the three-length-scale model is still lower than the one-length-scale model in which the estimation is qualitatively wrong.

### B. Target pattern synthesised by a function: two-dimensional dodecagonal quasi-crystal

The same algorithm was used for the target pattern synthesised by functions by equation (14) for the two-dimensional stripe, hexagonal, and dodecagonal quasi-crystal patterns, and the three-dimensional double-gyroid pattern. The results are summarised in Fig. 4 in the main text.

The target pattern of a dodecagonal quasi-crystal is synthesised by the function of superposition of twelve plane waves as equation (14) (see Methods in the main text). The pattern has twelve-fold rotational symmetry, as demonstrated in the pattern in the Fourier space (Fig. S5). We may generate a similar pattern from the estimated PDE to the target pattern, if not exactly the same.

The obtained histogram of the cost function is qualitatively similar to the target pattern of the numerically produced dodecagonal quasi-crystal (see Fig. S5). The two-length-scale model has two distinct energy scales associated with quasi-crystal and hexagonal patterns. At the lower cost function, we obtain quasi-crystals. The three-length-scale model also reproduces quasi-crystals, but the cost function is higher than that of the two-length-scale model. Therefore, the two-length-scale model will most likely give the target pattern.

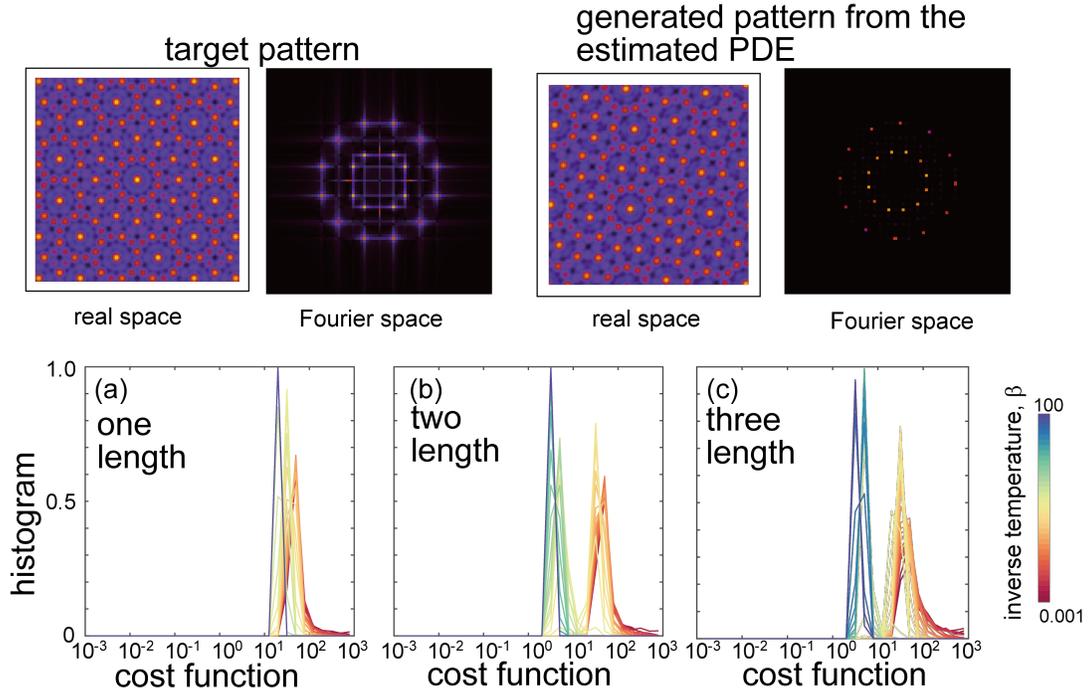


FIG. S5. (Top) The target pattern and the generated pattern from the estimated PDE in real and Fourier space. (Bottom) Histograms of the cost function of the REMC sampling for the target pattern of dodecagonal quasi-crystal that is synthesised by the function of equation (14). The horizontal axis is shown in the logarithmic scale.

The log marginal likelihood (free energy) at each temperature,  $\beta$ , is calculated for each model (see Fig. 5(c)). The models with one- and three-length scales show qualitatively the same  $\beta$  dependence of the log marginal likelihood with the results for numerically produced target pattern (Fig.3(e)). On the other hand, the two-length scale model  $m = m_2$  shows qualitatively different  $\beta$  dependence, namely there is a minimum at  $\beta \approx 16$ . This is because there is no ground truth in this target pattern. Intuitively, this situation is similar to the target pattern with noise. In both cases, the model of equation (5) cannot reproduce exactly the same pattern as the target. The Bayesian modelling is suitable in such cases because it gives us an optimal noise level[148]. Even though the cost function, or energy, is lower at the lower temperature,  $\beta$ , the posterior distribution is too narrow so that a slightly higher  $\beta$  than the lowest  $\beta$  gives the best estimate. This is exactly how our method estimate the observation noise in the target pattern, and avoids overfitting. When the problem does not have the ground truth, there is a nonzero observation error in the target pattern. By estimating the optimal  $\beta$ , we can quantify how the best-estimated model is close to the target pattern.

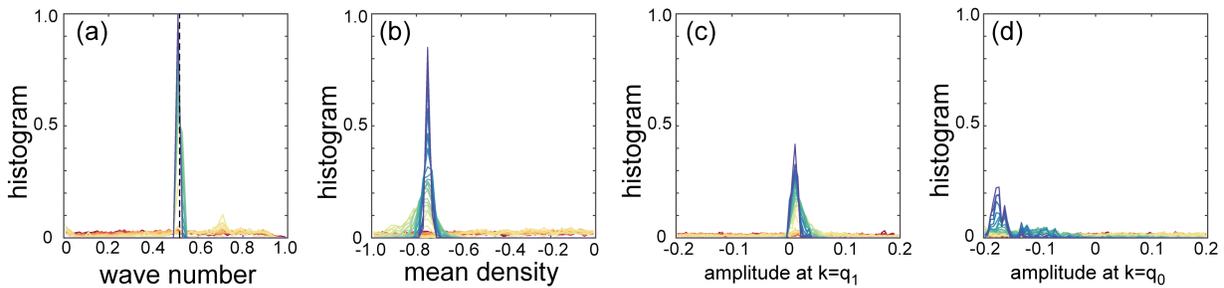


FIG. S6. Histogram of the estimated parameters for the target pattern of dodecagonal quasi-crystal that is synthesised by the function of equation (14). The vertical dashed line in the histogram of wavenumber indicates  $q_1 = \sqrt{2 - \sqrt{3}}$ , which is necessary to reproduce the dodecagonal pattern.

The estimated parameters are shown in Table S2, and the corresponding histograms for each  $\beta$  are shown in Fig. S6. The histograms of the sampled parameters are qualitatively similar to the results of the numerically produced target pattern. The distribution of the estimated wavenumber is well accumulated around the value that is necessary to make quasi-crystals, that is,  $q_1 \approx 0.51$ , which is close to  $1/(2 \cos(\pi/12))$  under  $q_0 = 1$ . The width of the distribution is sharp because if the ratio  $q_1/q_0$  does not satisfy the appropriate value, the coupling between the two modes cannot occur. Other parameters do not have specific values that they should satisfy, and therefore, show broader distribution than that of the wavelength. In fact, there is a region where the dodecagonal quasi-crystal can appear in the parameter space spanned by mean density  $\bar{\psi}$  and the spectrum amplitude  $a_1$ . These parameters are distributed inside the region. In particular, the mean density is a good quantity separating the dodecagonal quasi-crystal from the stripe and hexagonal patterns. This is because  $\bar{\psi}$  manifests the strength of coupling among three modes, that is, the coupling of two modes affects another mode in the dynamic equation (see the quadratic term in equation (9)). The stripe pattern has a symmetry of  $\psi \rightarrow -\psi$ , and therefore, this coupling must be weak. The hexagonal pattern does not have this symmetry, and this requires stronger coupling; the dodecagonal quasi-crystal requires even far stronger coupling[117]. Therefore,  $|\bar{\psi}| \gg 0$ . This is in agreement with the result of the estimation in Fig. S6(b). The distribution of  $a_0$  is broad for the same reason in the case of numerically produced dodecagonal quasi-crystal pattern.

TABLE S2. The estimated parameters values for functionally synthesised dodecagonal quasi-crystals. Errors are evaluated from the standard deviation of each sample from the REMC at the optimal temperature.

parameters	estimated (one length)	estimated (two length)	estimated (three length)
$dx = dy$	$0.986 \pm 0.0092$	$0.980 \pm 0.0018$	$0.992 \pm 0.0039$
$\bar{\psi}$	$-0.183 \pm 0.0497$	$-0.748 \pm 0.0192$	$-0.613 \pm 0.0149$
$a_0$	$0.010 \pm 0.0120$	$-0.132 \pm 0.0443$	$-0.0581 \pm 0.0213$
$q_1$	-	$0.514 \pm 0.0037$	$0.916 \pm 0.0174$
$a_1$	-	$0.0153 \pm 0.0085$	$-0.155 \pm 0.0214$
$q_2$	-	-	$0.507 \pm 0.0029$
$a_2$	-	-	$0.0245 \pm 0.00762$

### C. Target pattern synthesised by a function: stripe and hexagonal patterns

We apply our method to stripe and hexagonal patterns. These patterns are simpler than the dodecagonal quasi-crystal, and the one length scale is enough to reproduce the patterns. The histogram of the cost function is shown in Fig. S7 for the stripe pattern and in Fig. S9 for the hexagonal pattern. In contrast with the dodecagonal quasi-crystals, the cost function of these patterns can be very low up to about  $E \approx 10^{-5} - 10^{-4}$ . In addition, the shapes of the histogram under different models are qualitatively similar. Accordingly, the minimum log marginal likelihood and probability of each model are comparable among different models. Nevertheless, we estimate the one-length-scale model  $m = m_1$  is the best model in both stripe and hexagonal patterns (Table 4(a,b)). The dependence of the log marginal likelihood on  $\beta$  is indistinguishable among the three models. The marginal likelihood monotonically decreases as the temperature decreases (or  $\beta$  increases), and therefore, the minimum log marginal likelihood is attained at the

highest  $\beta$  (lowest temperature). The difference of the log marginal likelihood among the three models is small, but still, by looking at marginalised probability, we find the one-length-scale model is more likely than other models as shown in Table 4(a,b).

The generated patterns from the estimated PDE are similar to the target pattern in both stripe (Fig. S7) and hexagonal patterns (Fig. S9). We stress that thanks to the order parameter  $\Psi$ , we are able to identify two patterns up to translation and rotation. In fact, the generated pattern from the estimated parameters in Fig. S7 has a different orientation.

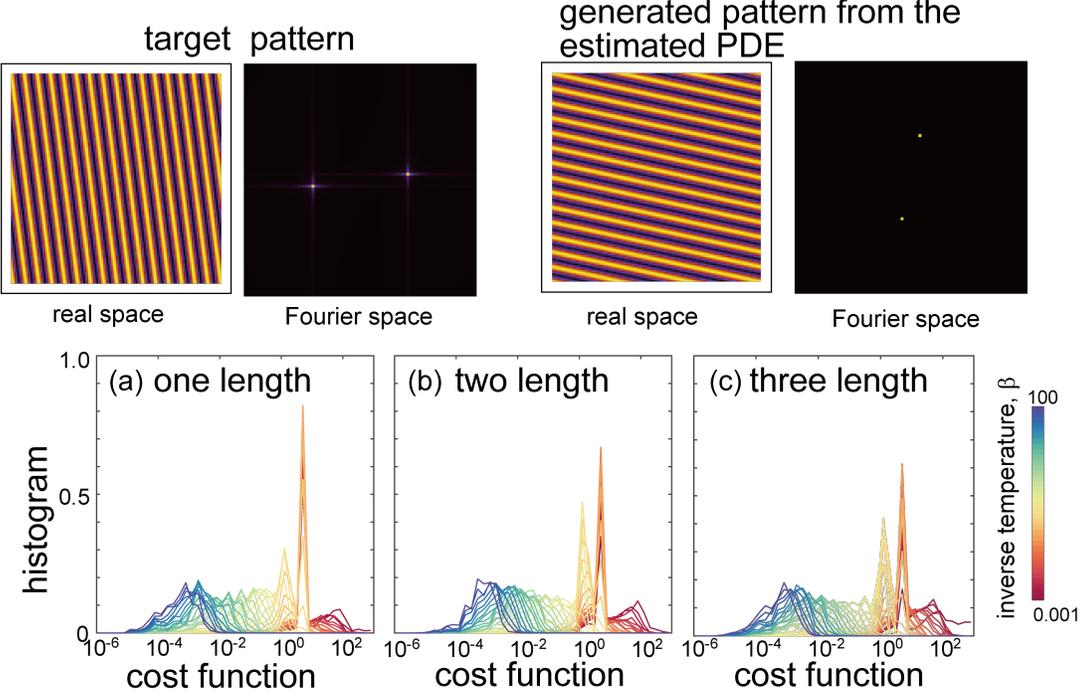


FIG. S7. (Top) The target pattern and the generated pattern from the estimated PDE in real and Fourier space. (Bottom) Histograms of the cost function of the REMC sampling for the target pattern of a stripe pattern that is synthesised by the function of equation (14). The horizontal axis is shown in the logarithmic scale.

The estimated parameters under the one-length-scale model are shown in Fig. S8 for the stripe pattern and in Fig. S10 for the hexagonal pattern. The estimated wavenumber is close to that of the target pattern  $q_0 = 0.5$  in both patterns. The relevant parameter to distinguish between the stripe and hexagonal patterns is the mean density  $\bar{\psi}$ . It is known that the stripe pattern appears near  $\bar{\psi} = 0$  because in this region, the pattern has parity symmetry,  $\psi = -\psi$  [111]. The hexagonal pattern breaks this symmetry, and also it requires the three-body interaction, which is the quadratic term in  $\psi$  in the dynamical equation (5). The hexagonal pattern is expressed by the superposition of three plain waves  $\psi_1 = e^{ix}$  and  $\psi_{2,3} = e^{i(-\frac{1}{2}x \pm \frac{\sqrt{3}}{2}y)}$ . The coupling between  $\psi_2$  and  $\psi_3$  gives rise to  $e^{-ix}$ , which is the complex conjugate to  $\psi_1$ , and thus this quadratic term appears in the equation of  $\psi_1$ . In fact, the stripe pattern appears near  $\bar{\psi} = 0$ , whereas the hexagonal pattern appears  $\psi \gg 0$ .

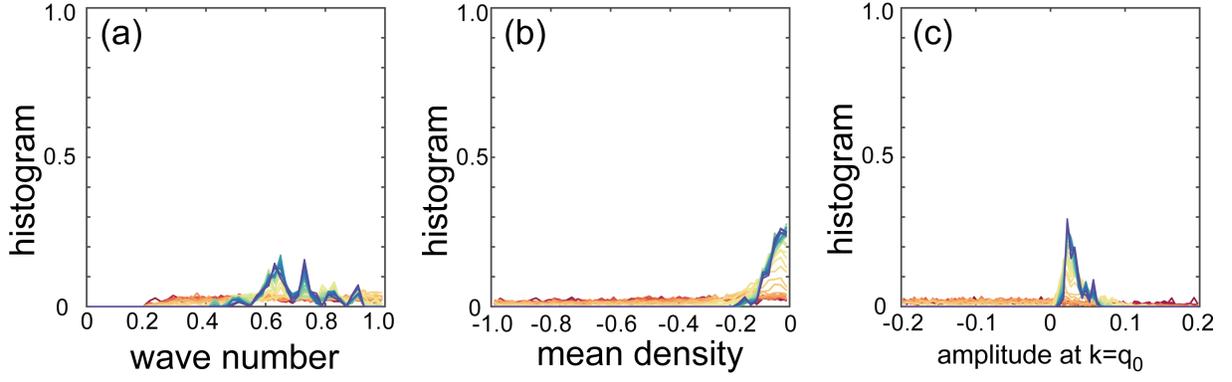


FIG. S8. Histogram of the estimated parameters for the target pattern of a stripe pattern that is synthesised by the function of equation (14).

TABLE S3. The estimated parameters values for functionally synthesised stripe pattern. Errors are evaluated from the standard deviation of each sample from the REMC at the optimal temperature.

parameters	estimated (one length)	estimated (two length)	estimated (three length)
$dx = dy$	$0.992 \pm 0.0153$	$0.995 \pm 0.0119$	$0.988 \pm 0.0146$
$\bar{\psi}$	$-0.0471 \pm 0.0351$	$-0.0648 \pm 0.0426$	$-0.0390 \pm 0.0281$
$a_0$	$0.0329 \pm 0.0115$	$-0.0337 \pm 0.0952$	$-0.0877 \pm 0.0788$
$q_1$	-	$0.571 \pm 0.127$	$0.728 \pm 0.240$
$a_1$	-	$0.0112 \pm 0.0420$	$-0.0781 \pm 0.0756$
$q_2$	-	-	$0.636 \pm 0.187$
$a_2$	-	-	$0.0097 \pm 0.0451$

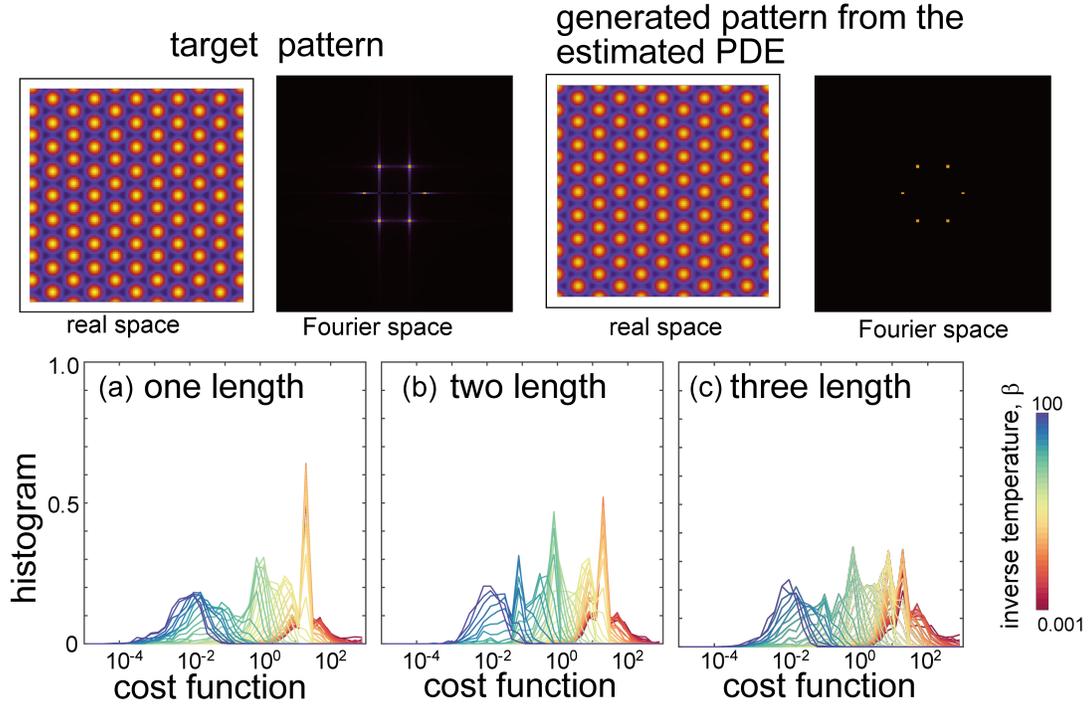


FIG. S9. (Top) The target pattern and the generated pattern from the estimated PDE in real and Fourier space. (Bottom) Histograms of the cost function of the REMC sampling for the target pattern of a hexagonal pattern that is synthesised by the function of equation (14). The horizontal axis is shown in the logarithmic scale.

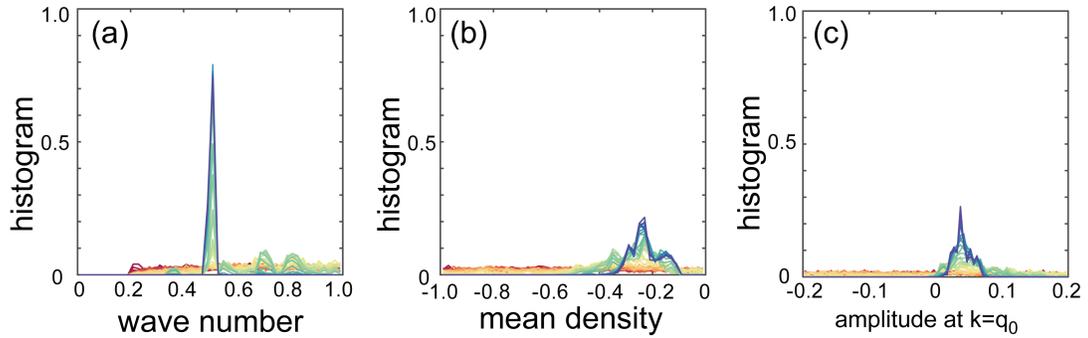


FIG. S10. Histogram of the estimated parameters for the target pattern of a hexagonal pattern that is synthesised by the function of equation (14).

TABLE S4. The estimated parameters values for a functionally synthesised hexagonal pattern. Errors are evaluated from the standard deviation of each data.

parameters	estimated (one length)	estimated (two length)	estimated (three length)
$dx = dy$	$0.990 \pm 0.00778$	$1.007 \pm 0.00642$	$1.03 \pm 0.00518$
$\bar{\psi}$	$-0.226 \pm 0.0536$	$-0.171 \pm 0.0445$	$-0.170 \pm 0.0399$
$a_0$	$0.0408 \pm 0.0126$	$0.0438 \pm 0.0341$	$-0.121 \pm 0.0450$
$q_1$	-	$0.492 \pm 0.0162$	$0.635 \pm 0.113$
$a_1$	-	$0.0534 \pm 0.00713$	$-0.102 \pm 0.0408$
$q_2$	-	-	$0.483 \pm 0.00261$
$a_2$	-	-	$0.0424 \pm 0.00586$

#### D. Target pattern synthesised by a function: double gyroid and Frank Kasper A15

Estimation of a PDE that generates a three-dimensional pattern is performed similarly to a two-dimensional pattern. The only difference is the definition of order parameter  $\Psi$ . As discussed in Methods, three-dimensional patterns may have several invariants for each  $l$ . We expect including all the invariants is necessary to classify complex patterns, but in this study, we consider only one invariant similar to the two-dimensional problem. We found this assumption works at least for the patterns that we study here: double gyroid (DG) and Frank Kasper A15 (FKA15). We have confirmed our method also works for a simpler pattern such as lamellae pattern (stripe in three dimensions) and cylindrical pattern (hexagonal pattern in three dimensions).

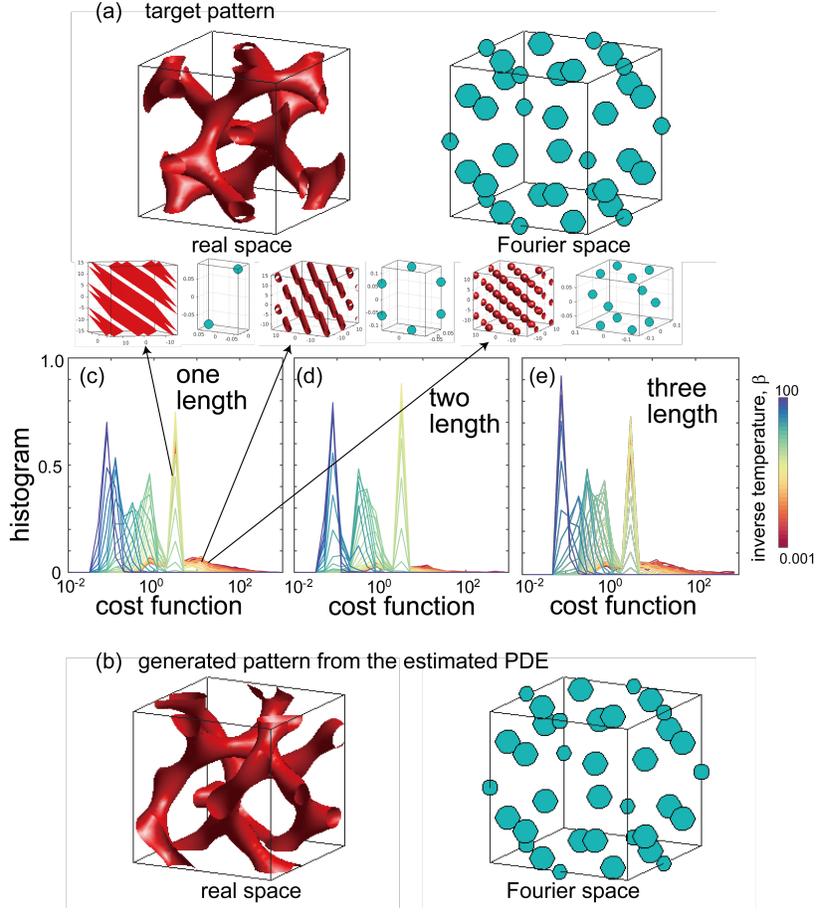


FIG. S11. (a,b) The target pattern and the generated pattern from the estimated PDE in real  $\psi(\mathbf{x})$  and Fourier  $|\tilde{\psi}(\mathbf{k})|$  space for the target pattern of double gyroid. In Fourier space, each point corresponds to a peak of  $|\tilde{\psi}(\mathbf{k})|$ , and the intensity is described by the size of the points. (c-e) Histogram of the cost function of the REMC sampling for the target pattern of a double gyroid pattern that is synthesised by the function as equation (14). The horizontal axis is shown in the logarithmic scale. Patterns with higher energy are also shown in the insets.

In Fig. S11, the histogram of the cost function of each model is shown. The DG pattern is reproduced by all models ( $m = m_1, m_2, m_3$ ). In fact, the generated pattern from the estimated PDE is identical to the target pattern up to translation, and they have almost the same patterns in the Fourier space (Fig. S11). The histograms of the cost function of all models are qualitatively similar, but  $m = m_2$  has slightly smaller energy on average. The probability of the two models shown in Fig. 4(d) is comparable, but  $m = m_2$  is chosen. The log marginal likelihood decreases as the temperature decreases ( $\beta$  increases), and therefore, within the current choice of the range of  $\beta$ , the minimum of the log marginal likelihood is not attained. If we use the temperature range containing larger  $\beta$ , the log marginal likelihood may attain its minimum.

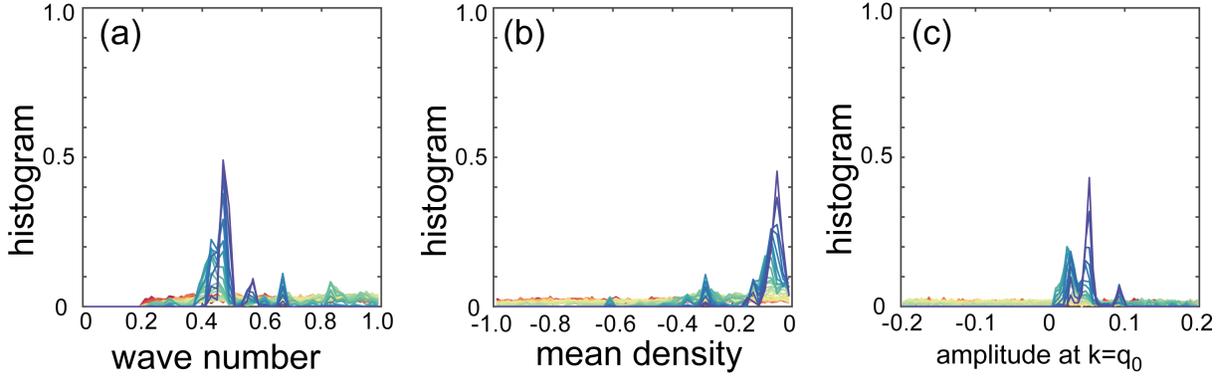


FIG. S12. Histogram of the estimated parameters for the target pattern of a double gyroid pattern that is synthesised by the function of equation (14).

Estimated parameters are shown in Fig. S12 and Table. S5. The estimated wavenumber distributes near the wavelength that we have imposed. The mean density  $\bar{\psi}$  is accumulated near  $\bar{\psi} = 0$  but with a slight deviation. This is consistent with previous theoretical and numerical results; the double gyroid pattern, in fact, appears between lamellae (stripe) and cylinder (hexagonal) patterns[107, 109, 190]. As discussed in Sec. SIV B and Sec. SIV C, the stripe pattern appears near  $\bar{\psi} = 0$  including  $\bar{\psi} = 0$  whereas the hexagonal pattern appears at  $|\bar{\psi}| \gg 0$ . The region where the double gyroid pattern appears is between the two regions, and thus the mean density should be  $1 \gg |\bar{\psi}| > 0$ . This is exactly observed in Fig.S12(b).

TABLE S5. The estimated parameters values for functionally synthesised double gyroid pattern in three dimensions. Errors are evaluated from the standard deviation of each sample from the REMC at the optimal temperature.

parameters	estimated (one length)	estimated (two length)	estimated (three length)
$dx = dy = dz$	$0.528 \pm 0.0261$	$0.561 \pm 0.00380$	$0.565 \pm 0.00995$
$\bar{\psi}$	$-0.0559 \pm 0.0326$	$-0.090 \pm 0.0266$	$-0.165 \pm 0.0102$
$a_0$	$0.0548 \pm 0.0149$	$0.127 \pm 0.0236$	$0.0713 \pm 0.0123$
$q_1$	-	$0.838 \pm 0.0187$	$0.349 \pm 0.188$
$a_1$	-	$0.112 \pm 0.0208$	$-0.0604 \pm 0.0336$
$q_2$	-	-	$0.855 \pm 0.0156$
$a_2$	-	-	$0.146 \pm 0.0398$

The FKA15 pattern is expressed by 24 wave vectors  $\mathbf{q}^* = (\pm 2, \pm 1, 0)$ , 24 wave vectors  $\mathbf{q}^* = (\pm 2, \pm 1, \pm 1)$ , 6 wave vectors of  $\mathbf{q}^* = (\pm 2, 0, 0)$  with their permutation along the  $x, y, z$  directions[191]. In the unit cell, a centre particle is surrounded by 8 particles at the position of the corners, and 2 particles are located at each face (Fig. S13(a)). In Fig. S13, the histogram of the cost function for each model is shown. The FKA15 pattern is reproduced by all models ( $m = m_1, m_2, m_3$ ). The generated pattern from the estimated PDE is similar to the target pattern up to translation, and they have almost the same patterns in the Fourier space (Fig. S13(b)). The similarity is also demonstrated in their side views. The histograms of the cost function are shown in Fig. S13(c-e). The probability of the two models shown in Fig. 4(e) is comparable, but  $m = m_1$  is chosen for this target pattern. The log marginal likelihood decreases as the temperature decreases, and therefore, within the current choice of the range of temperature, the minimum of the log marginal likelihood is not attained. If we use the temperature range containing larger  $\beta$ , the log marginal likelihood may attain its minimum.

Estimated parameters are shown in Fig. S13(f-h). The estimated wavenumber distributes near the wavelength that we have imposed  $q = 0.5$ . In contrast with the DG pattern, the mean density  $\bar{\psi}$  is away from  $\bar{\psi} = 0$ . In fact, FKA15 has been found in higher resolution of  $\bar{\psi} - a_0$  phase diagram using the self-consistent field theory, which describes block copolymers[192]. The position of FKA15 in the phase diagram used to be BCC, but recently several Frank Kasper phases have been found in this region. The BCC pattern has been found at  $\bar{\psi} \gg 0$ , and thus our estimated  $\bar{\psi}$  is consistent with the observation. We note that, to our knowledge, FKA15 has not been reported in the framework

of PFC. Because PFC is considered as an approximation of various model, including self-consistent field theory and density functional theory of atomic alloy, this pattern should appear in various pattern-forming systems.

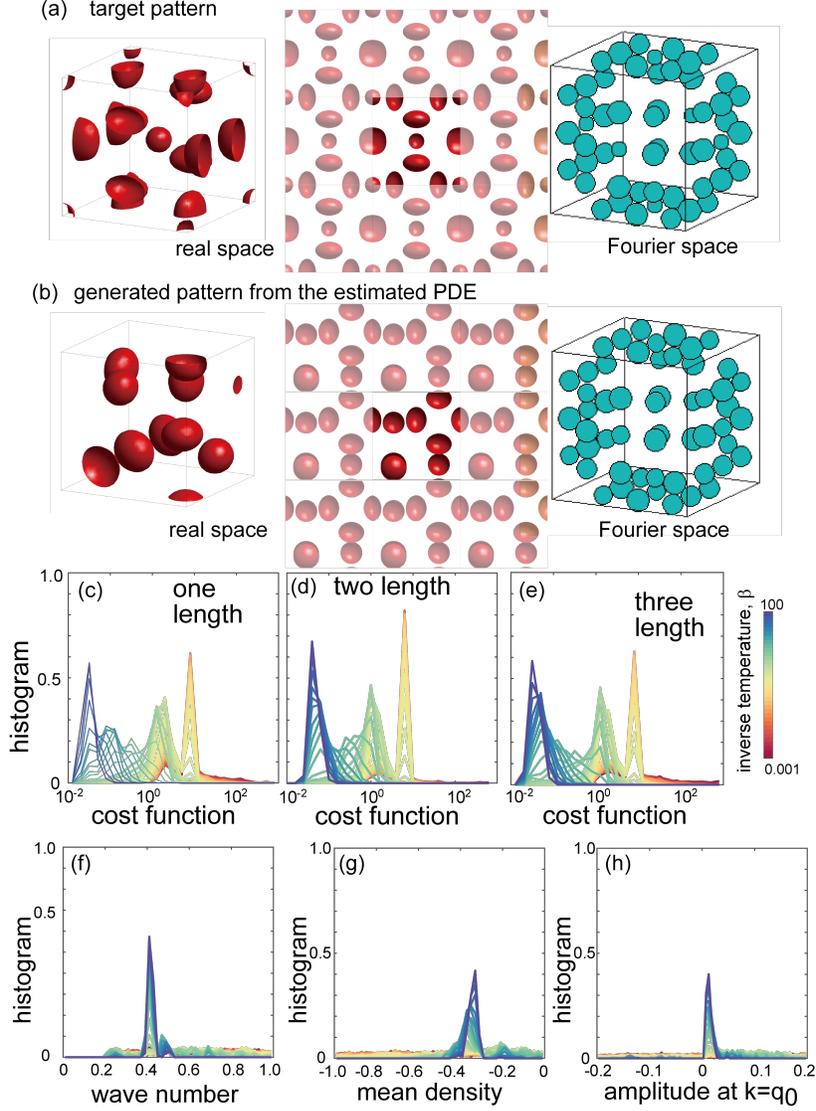


FIG. S13. (a,b) The target (a) and estimated (b) patterns in real  $\psi(\mathbf{x})$  and Fourier  $|\psi(\mathbf{k})|$  space for the target pattern of Frank Kasper A15. In Fourier space, each point corresponds to a peak of  $|\psi(\mathbf{k})|$ , and the intensity is described by the size of the points. The side view of the real-space images is also shown in the middle panels. In the side view, neighbouring domains of the periodic boundary are added around the main domain to clarify periodicity. (c-e) Histogram of the cost function in the REMC sampling for the target pattern of a Frank Kasper A15 pattern that is synthesised by the function of equation (14). The horizontal axis is shown in the logarithmic scale. Patterns with higher energy are also shown in the insets. (f-h) Histogram of the estimated parameters for the target pattern of a Frank Kasper A15 pattern that is synthesised by the function.

TABLE S6. The estimated parameters values for functionally synthesised Frank Kasper A15 pattern in three dimensions from experimental data. Errors are evaluated from the standard deviation of each sample from the REMC at the optimal temperature.

parameters	estimated (one length)	estimated (two length)	estimated (three length)
$dx = dy = dz$	$0.533 \pm 0.0364$	$0.461 \pm 0.0037$	$0.466 \pm 0.00705$
$\bar{\psi}$	$-0.341 \pm 0.0235$	$-0.335 \pm 0.0357$	$-0.384 \pm 0.0361$
$a_0$	$0.0131 \pm 0.005$	$0.101 \pm 0.0483$	$0.118 \pm 0.0377$
$q_1$	-	$0.889 \pm 0.0270$	$0.919 \pm 0.0264$
$a_1$	-	$0.0673 \pm 0.0721$	$0.0532 \pm 0.0267$
$q_2$	-	-	$0.599 \pm 0.0533$
$a_2$	-	-	$-0.140 \pm 0.0473$

Each domain of the target pattern of FK A15 is deformed (Fig. S13(a)). Figure S13(b) shows that the generated patterns using the estimated parameters are also deformed. To see the deviation from a spherical shape quantitatively, we extract the centre of each domain in the density field  $\psi(\mathbf{x})$  of the generated patterns. The density field is binarised using the threshold of half of the maximum density. Then, the connected components are extracted from the binary field, and the centre of each connected component is computed. We also compute a gyration tensor of each domain, from which three eigenvalues associated with lengths of three axes of an ellipsoid. Sphericity is defined by the ratio between the maximum and minimum eigenvalues. When the domain is a spherical shape, sphericity is one, whereas it is larger than one when the shape is deformed. Figure S14 shows the correlation between sphericity and the cost function. The one-length-scale model  $m_1$  generates patterns that are more deformed than those generated by the two-length-scale model  $m_2$ . Accordingly, the cost functions of the generated patterns by the one-length-scale model are lower.

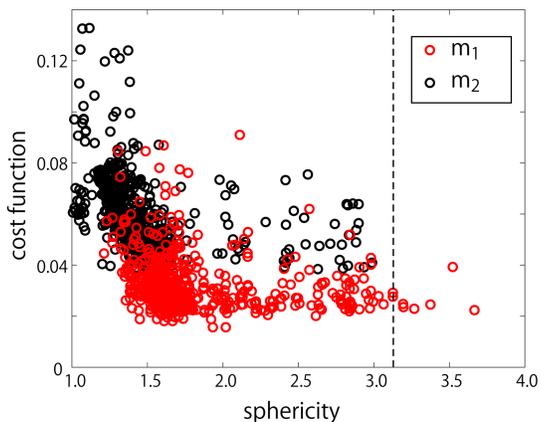


FIG. S14. The correlation between sphericity and the cost function for the generated patterns using the estimated parameters for one-length-scale model  $m_1$  and two-length-scale model  $m_2$ . The sphericity of the target pattern is shown in a vertical dashed line.

## SV. DAMAGED TARGET PATTERN

In addition to the estimation for the target pattern with noise in the text, we perform the estimation for the damaged target pattern. The damaged pattern is made by setting  $\psi = 0$  at randomly distributed spatial points. We vary the fraction of the damaged point from 0 to 50 %, and estimate the wave number, which is most relevant parameter to reproduce dodecagonal QC. Figure S15 shows that up to 30 % of the damage, the correct wave number can be estimated.

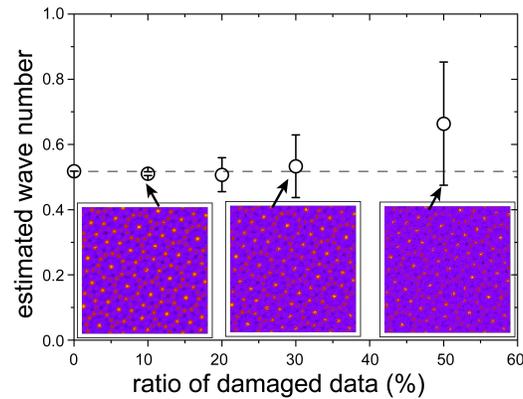


FIG. S15. Estimated wavenumbers for numerically generated QC by BM-PDE equation (2) for the damaged target pattern.

- 
- [1] Swift, J. & Hohenberg, P. C. Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A* **15**, 319–328 (1977).
  - [2] Lord Rayleigh, O. F. Lix. on convection currents in a horizontal layer of fluid, when the higher temperature is on the under side. *Lond. Edinb. Dubl. Phil. Mag.* **32**, 529–546 (1916).
  - [3] Taylor, G. I. Viii. stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. Royal Soc. A* **223**, 289–343 (1923).
  - [4] Stuart, J. T. On the non-linear mechanics of hydrodynamic stability. *J. Fluid Mech.* **4**, 1–21 (1958).
  - [5] Watson, J. On the non-linear mechanics of wave disturbances in stable and unstable parallel flows part 2. the development of a solution for plane poiseuille flow and for plane couette flow. *J. Fluid Mech.* **9**, 371–389 (1960).
  - [6] Lega, J., Moloney, J. V. & Newell, A. C. Swift-hohenberg equation for lasers. *Phys. Rev. Lett.* **73**, 2978–2981 (1994).
  - [7] Lefever, R., Barbier, N., Couteron, P. & Lejeune, O. Deeply gapped vegetation patterns: On crown/root allometry, criticality and desertification. *J. Theor. Biol.* **261**, 194 – 209 (2009).
  - [8] Ohta, T. & Kawasaki, K. Equilibrium morphology of block copolymer melts. *Macromolecules* **19**, 2621–2632 (1986).
  - [9] Bahiana, M. & Oono, Y. Cell dynamical system approach to block copolymers. *Phys. Rev. A* **41**, 6763 (1990).
  - [10] Fredrickson, G. H. & Helfand, E. Fluctuation effects in the theory of microphase separation in block copolymers. *J. Chem. Phys.* **87**, 697–705 (1987).
  - [11] Podneks, V. E. & Hamley, I. W. Landau- brazovskii theory for the ia3d structure. *J. Exp. Theor. Phys. Lett.* **64**, 617–624 (1996).
  - [12] Pismen, L. M. *Patterns and interfaces in dissipative dynamics* (Springer, 2006).
  - [13] Jaatinen, A. & Ala-Nissila, T. Extended phase diagram of the three-dimensional phase field crystal model. *J. Phys. Cond. Mat.* **22**, 205402 (2010).
  - [14] Nonomura, M. & Ohta, T. Kinetics of morphological transitions between mesophases. *J. Phys. Cond. Mat.* **13**, 9089–9112 (2001).
  - [15] Elder, K. R., Katakowski, M., Haataja, M. & Grant, M. Modeling elasticity in crystal growth. *Phys. Rev. Lett.* **88**, 245701 (2002).
  - [16] Provatas, N. & Elder, K. *Phase-field methods in materials science and engineering* (John Wiley & Sons, 2011).
  - [17] Kléman, M. *Points, Lines and Walls: in Liquid Crystals, Magnetic Systems, and Various Disordered Media* (Wiley, New York, 1983).
  - [18] Harrison, C. *et al.* Mechanisms of ordering in striped patterns. *Science* **290**, 1558–1560 (2000).
  - [19] Kléman, M. & Laverntovich, O. D. *Soft matter physics: an introduction* (Springer, 2003).
  - [20] Elder, K. R., Provatas, N., Berry, J., Stefanovic, P. & Grant, M. Phase-field crystal modeling and classical density functional theory of freezing. *Phys. Rev. B* **75**, 064107 (2007).
  - [21] Lifshitz, R. & Petrich, D. M. Theoretical model for faraday waves with multiple-frequency forcing. *Phys. Rev. Lett.* **79**, 1261–1264 (1997).
  - [22] Savitz, S., Babadi, M. & Lifshitz, R. Multiple-scale structures: from Faraday waves to soft-matter quasicrystals. *IUCrJ* **5**, 247–268 (2018).
  - [23] Subramanian, P., Archer, A. J., Knobloch, E. & Rucklidge, A. M. Three-dimensional icosahedral phase field quasicrystal. *Phys. Rev. Lett.* **117**, 075501 (2016).
  - [24] Müller, H. W. Model equations for two-dimensional quasipatterns. *Phys. Rev. E* **49**, 1273–1277 (1994).
  - [25] Mermin, N. & Troian, S. M. Mean-field theory of quasicrystalline order. *Phys. Rev. Lett.* **54**, 1524 (1985).
  - [26] Kats, E., Lebedev, V. & Muratov, A. Weak crystallization theory. *Phys. Rep.* **228**, 1–91 (1993).

- [27] Malomed, B., Nepomnyashchii, A. & Tribelskii, M. Two-dimensional quasiperiodic structures in nonequilibrium systems. *Sov. Phys. JETP* (1988).
- [28] Zhang, Z. & Glotzer, C., Sharon. Self-assembly of patchy particles. *Nano Lett.* **4**, 1407–1413 (2004).
- [29] Damasceno, P. F., Engel, M. & Glotzer, S. C. Predictive self-assembly of polyhedra into complex structures. *Science* **337**, 453–457 (2012).
- [30] Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418–2421 (2002).
- [31] Li, F., Josephson, D. P. & Stein, A. Colloidal assembly: The road from particles to colloidal molecules and crystals. *Angew. Chem. Int. Edit.* **50**, 360–388 (2011).
- [32] Boles, M. A., Engel, M. & Talapin, D. V. Self-assembly of colloidal nanocrystals: From intricate structures to functional materials. *Chem. Rev.* **116**, 11220–11289 (2016). PMID: 27552640.
- [33] Rechtsman, M. C., Stillinger, F. H. & Torquato, S. Optimized interactions for targeted self-assembly: Application to a honeycomb lattice. *Phys. Rev. Lett.* **95**, 228301 (2005).
- [34] Rechtsman, M., Stillinger, F. & Torquato, S. Designed interaction potentials via inverse methods for self-assembly. *Phys. Rev. E* **73**, 011406 (2006).
- [35] Torquato, S. Inverse optimization techniques for targeted self-assembly. *Soft Matter* **5**, 1157–1173 (2009).
- [36] Torquato, S. Optimal design of heterogeneous materials. *Annu. Rev. Mater. Res.* **40**, 101–129 (2010).
- [37] Jain, A., Bollinger, J. A. & Truskett, T. M. Inverse methods for material design. *AIChE J.* **60**, 2732–2740 (2014).
- [38] Sherman, Z. M., Howard, M. P., Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Inverse methods for design of soft materials. *J. Chem. Phys.* **152**, 140902 (2020).
- [39] Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Communication: Inverse design for self-assembly via on-the-fly optimization. *J. Chem. Phys.* **145**, 111101 (2016).
- [40] Tsai, C. L., Delaney, K. T. & Fredrickson, G. H. Genetic algorithm for discovery of globally stable phases in block copolymers. *Macromolecules* **49**, 6558–6567 (2016).
- [41] Paradiso, S. P., Delaney, K. T. & Fredrickson, G. H. Swarm intelligence platform for multiblock polymer inverse formulation design. *ACS Macro Letters* **5**, 972–976 (2016).
- [42] Khadilkar, M. R., Paradiso, S., Delaney, K. T. & Fredrickson, G. H. Inverse design of bulk morphologies in multiblock polymers using particle swarm optimization. *Macromolecules* **50**, 6702–6709 (2017).
- [43] Khaira, G. S. *et al.* Evolutionary optimization of directed self-assembly of triblock copolymers on chemically patterned substrates. *ACS Macro Lett.* **3**, 747–752 (2014).
- [44] Kumar, R., Coli, G. M., Dijkstra, M. & Sastry, S. Inverse design of charged colloidal particle interactions for self assembly into specified crystal structures. *J. Chem. Phys.* **151**, 084109 (2019).
- [45] Miskin, M. Z., Khaira, G., de Pablo, J. J. & Jaeger, H. M. Turning statistical physics models into materials design engines. *Proc. Nat. Acad. Sci. USA* **113**, 34–39 (2016).
- [46] Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Assembly of nothing: equilibrium fluids with designed structured porosity. *Soft Matter* **12**, 2663–2667 (2016).
- [47] Piñeros, W. D., Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Inverse design of multicomponent assemblies. *J. Chem. Phys.* **148**, 104509 (2018).
- [48] Piñeros, W. D., Jadrlich, R. B. & Truskett, T. M. Design of two-dimensional particle assemblies using isotropic pair interactions with an attractive well. *AIP Adv.* **7**, 115307 (2017).
- [49] Adorf, C. S., Antonaglia, J., Dshemuchadse, J. & Glotzer, S. C. Inverse design of simple pair potentials for the self-assembly of complex structures. *J. Chem. Phys.* **149**, 204102 (2018).
- [50] Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Communication: From close-packed to topologically close-packed: Formation of laves phases in moderately polydisperse hard-sphere mixtures. *J. Chem. Phys.* **148**, 191101 (2018).
- [51] Jaeger, H. M. & de Pablo, J. J. Perspective: Evolutionary design of granular media and block copolymer patterns. *APL Mater.* **4**, 053209 (2016).
- [52] Tokuda, S., Nagata, K. & Okada, M. Simultaneous estimation of noise variance and number of peaks in bayesian spectral deconvolution. *J. Phys. Soc. Japan* **86**, 024001 (2017).
- [53] Bär, M., Hegger, R. & Kantz, H. Fitting partial differential equations to space-time dynamics. *Phys. Rev. E* **59**, 337–342 (1999).
- [54] Müller, T. G. & Timmer, J. Parameter identification techniques for partial differential equations. *Int. J. Bifurcat. Chaos* **14**, 2053–2060 (2004).
- [55] Zhao, H., Storey, B. D., Braatz, R. D. & Bazant, M. Z. Learning the physics of pattern formation from images. *Phys. Rev. Lett.* **124**, 060201 (2020).
- [56] Voss, H. U., Timmer, J. & Kurths, J. Nonlinear dynamical system identification from uncertain and indirect measurements. *Int. J. Bifurcat. Chaos* **14**, 1905–1933 (2004).
- [57] Bongard, J. & Lipson, H. Automated reverse engineering of nonlinear dynamical systems. *Proc. Nat. Acad. Sci. USA* **104**, 9943–9948 (2007).
- [58] Schmidt, M. & Lipson, H. Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).
- [59] Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Nat. Acad. Sci. USA* **113**, 3932–3937 (2016).
- [60] Schaeffer, H. Learning partial differential equations via data discovery and sparse optimization. *Proc. Royal Soc. A* **473**, 20160446 (2017).
- [61] Rudy, S. H., Brunton, S. L., Proctor, J. L. & Kutz, J. N. Data-driven discovery of partial differential equations. *Sci. Adv.* **3**, e1602614 (2017).

- [62] Schaeffer, H. & McCalla, S. G. Sparse model selection via integral terms. *Phys. Rev. E* **96**, 023302 (2017).
- [63] Xu, H., Chang, H. & Zhang, D. DL-pde: Deep-learning based data-driven discovery of partial differential equations from discrete and noisy data. *arXiv:1908.04463* (2019).
- [64] Bishop, C. M. *Pattern recognition and machine learning I*, vol. 1 (Springer New York, 2006).
- [65] Kaipio, J. & Somersalo, E. *Statistical and computational inverse problems*, vol. 160 (Springer Science & Business Media, 2006).
- [66] Xun, X., Cao, J., Mallick, B., Maity, A. & Carroll, R. J. Parameter estimation of partial differential equation models. *J. Am. Stat. Assoc.* **108**, 1009–1020 (2013).
- [67] Campillo-Funollet, E., Venkataraman, C. & Madzvamuse, A. Bayesian parameter identification for Turing systems on stationary and evolving domains. *Bull. Math. Biol.* **81**, 81–104 (2019).
- [68] Kazarnikov, A. & Haario, H. Statistical approach for parameter identification by Turing patterns. *Journal of Theoretical Biology* **501**, 110319 (2020).
- [69] Zhao, H., Braatz, R. D. & Bazant, M. Z. Image inversion and uncertainty quantification for constitutive laws of pattern formation. *J. Comp. Phys.* **436**, 110279 (2021).
- [70] Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6**, 187–202 (2009).
- [71] Law, K., Stuart, A. & Zygalakis, K. *Data assimilation* (Springer, 2015).
- [72] Carrassi, A., Bocquet, M., Hannart, A. & Ghil, M. Estimating model evidence using data assimilation. *Q. J. R. Meteorol. Soc.* **143**, 866–880 (2017).
- [73] Evensen, G. *Data assimilation: the ensemble Kalman filter* (Springer, 2009).
- [74] Ito, S.-i., Nagao, H., Kurokawa, T., Kasuya, T. & Inoue, J. Bayesian inference of grain growth prediction via multi-phase-field models. *Phys. Rev. Mat.* **3**, 053404 (2019).
- [75] Yuan, Y. *et al.* Machine discovery of partial differential equations from spatiotemporal data. *arXiv:1909.06730* (2019).
- [76] Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005).
- [77] Tenorio, L. *An introduction to data analysis and uncertainty quantification for inverse problems*, vol. 3 (SIAM, 2017).
- [78] Varah, J. M. A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing* **3**, 28–46 (1982).
- [79] Ramsay, J. O., Hooker, G., Campbell, D. & Cao, J. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 741–796 (2007).
- [80] Rasmussen, C. E. *Gaussian processes in machine learning* (2003).
- [81] Graepel, T. Solving noisy linear operator equations by Gaussian processes: Application to ordinary and partial differential equations. In *ICML*, 234–241 (2003).
- [82] Álvarez, M., Luengo, D. & Lawrence, N. D. Latent force models. vol. 5 of *Proceedings of Machine Learning Research*, 9–16 (PMLR, 2009).
- [83] Melkumyan, A. Operator induced multi-task Gaussian processes for solving differential equations. In *Neural Information Processing Systems (NIPS) Workshop: New Directions in Multiple Kernel Learning* (2012).
- [84] Atkinson, S. & Zabaras, N. Structured Bayesian Gaussian process latent variable model: Applications to data-driven dimensionality reduction and high-dimensional inversion. *Journal of Computational Physics* **383**, 166–195 (2019).
- [85] Calderhead, B., Girolami, M. & Lawrence, N. D. Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes. In Koller, D., Schuurmans, D., Bengio, Y. & Bottou, L. (eds.) *Advances in Neural Information Processing Systems 21*, 217–224 (2009).
- [86] Dondelinger, F., Husmeier, D., Rogers, S. & Filippone, M. Ode parameter inference using adaptive gradient matching with Gaussian processes. vol. 31 of *Proceedings of Machine Learning Research*, 216–228 (PMLR, 2013).
- [87] Gorbach, N. S., Bauer, S. & Buhmann, J. M. Scalable variational inference for dynamical systems. In *Advances in Neural Information Processing Systems*, 4806–4815 (2017).
- [88] Raissi, M., Perdikaris, P. & Karniadakis, G. Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing* **40**, A172–A198 (2018).
- [89] Baake, E., Baake, M., Bock, H. G. & Briggs, K. M. Fitting ordinary differential equations to chaotic data. *Phys. Rev. A* **45**, 5524–5529 (1992).
- [90] Müller, T. & Timmer, J. Fitting parameters in partial differential equations from partially observed noisy data. *Physica D: Nonlinear Phenomena* **171**, 1–7 (2002).
- [91] Foucart, S. & Rauhut, H. *A Mathematical Introduction to Compressive Sensing* (Springer Science & Business Media, New York, NY, 2013).
- [92] Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. Royal Stat. Soc. B* **67**, 301–320 (2005).
- [93] Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
- [94] Shi, A.-C. Nature of anisotropic fluctuation modes in ordered systems. *J. Phys. Cond. Mat.* **11**, 10183 (1999).
- [95] Impéror-Clerc, M. Three-dimensional periodic complex structures in soft matter: investigation using scattering methods. *Interface focus* **2**, 589–601 (2012).
- [96] Bates, M. W. *et al.* Stability of the a15 phase in diblock copolymer melts. *Proc. Nat. Acad. Sci. USA* **116**, 13194–13199 (2019).
- [97] Swift, J. & Hohenberg, P. C. Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A* **15**, 319–328 (1977).
- [98] Lord Rayleigh, O. F. Lix. on convection currents in a horizontal layer of fluid, when the higher temperature is on the

- under side. *Lond. Edinb. Dubl. Phil. Mag.* **32**, 529–546 (1916).
- [99] Taylor, G. I. Viii. stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. Royal Soc. A* **223**, 289–343 (1923).
- [100] Stuart, J. T. On the non-linear mechanics of hydrodynamic stability. *J. Fluid Mech.* **4**, 1–21 (1958).
- [101] Watson, J. On the non-linear mechanics of wave disturbances in stable and unstable parallel flows part 2. the development of a solution for plane poiseuille flow and for plane couette flow. *J. Fluid Mech.* **9**, 371–389 (1960).
- [102] Lega, J., Moloney, J. V. & Newell, A. C. Swift-hohenberg equation for lasers. *Phys. Rev. Lett.* **73**, 2978–2981 (1994).
- [103] Lefever, R., Barbier, N., Couteron, P. & Lejeune, O. Deeply gapped vegetation patterns: On crown/root allometry, criticality and desertification. *J. Theor. Biol.* **261**, 194 – 209 (2009).
- [104] Ohta, T. & Kawasaki, K. Equilibrium morphology of block copolymer melts. *Macromolecules* **19**, 2621–2632 (1986).
- [105] Bahiana, M. & Oono, Y. Cell dynamical system approach to block copolymers. *Phys. Rev. A* **41**, 6763 (1990).
- [106] Fredrickson, G. H. & Helfand, E. Fluctuation effects in the theory of microphase separation in block copolymers. *J. Chem. Phys.* **87**, 697–705 (1987).
- [107] Podneps, V. E. & Hamley, I. W. Landau- brazovskii theory for the ia3d structure. *J. Exp. Theor. Phys. Lett.* **64**, 617–624 (1996).
- [108] Pismen, L. M. *Patterns and interfaces in dissipative dynamics* (Springer, 2006).
- [109] Jaatinen, A. & Ala-Nissila, T. Extended phase diagram of the three-dimensional phase field crystal model. *J. Phys. Cond. Mat.* **22**, 205402 (2010).
- [110] Nonomura, M. & Ohta, T. Kinetics of morphological transitions between mesophases. *J. Phys. Cond. Mat.* **13**, 9089–9112 (2001).
- [111] Elder, K. R., Katakowski, M., Haataja, M. & Grant, M. Modeling elasticity in crystal growth. *Phys. Rev. Lett.* **88**, 245701 (2002).
- [112] Provatas, N. & Elder, K. *Phase-field methods in materials science and engineering* (John Wiley & Sons, 2011).
- [113] Kléman, M. *Points, Lines and Walls: in Liquid Crystals, Magnetic Systems, and Various Disordered Media* (Wiley, New York, 1983).
- [114] Harrison, C. *et al.* Mechanisms of ordering in striped patterns. *Science* **290**, 1558–1560 (2000).
- [115] Kléman, M. & Laverntovich, O. D. *Soft matter physics: an introduction* (Springer, 2003).
- [116] Elder, K. R., Provatas, N., Berry, J., Stefanovic, P. & Grant, M. Phase-field crystal modeling and classical density functional theory of freezing. *Phys. Rev. B* **75**, 064107 (2007).
- [117] Lifshitz, R. & Petrich, D. M. Theoretical model for faraday waves with multiple-frequency forcing. *Phys. Rev. Lett.* **79**, 1261–1264 (1997).
- [118] Savitz, S., Babadi, M. & Lifshitz, R. Multiple-scale structures: from Faraday waves to soft-matter quasicrystals. *IUCrJ* **5**, 247–268 (2018).
- [119] Subramanian, P., Archer, A. J., Knobloch, E. & Rucklidge, A. M. Three-dimensional icosahedral phase field quasicrystal. *Phys. Rev. Lett.* **117**, 075501 (2016).
- [120] Müller, H. W. Model equations for two-dimensional quasipatterns. *Phys. Rev. E* **49**, 1273–1277 (1994).
- [121] Mermin, N. & Troian, S. M. Mean-field theory of quasicrystalline order. *Phys. Rev. Lett.* **54**, 1524 (1985).
- [122] Kats, E., Lebedev, V. & Muratov, A. Weak crystallization theory. *Phys. Rep.* **228**, 1–91 (1993).
- [123] Malomed, B., Nepomnyashchii, A. & Tribelskii, M. Two-dimensional quasiperiodic structures in nonequilibrium systems. *Sov. Phys. JETP* (1988).
- [124] Zhang, Z. & Glotzer, C., Sharon. Self-assembly of patchy particles. *Nano Lett.* **4**, 1407–1413 (2004).
- [125] Damasceno, P. F., Engel, M. & Glotzer, S. C. Predictive self-assembly of polyhedra into complex structures. *Science* **337**, 453–457 (2012).
- [126] Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418–2421 (2002).
- [127] Li, F., Josephson, D. P. & Stein, A. Colloidal assembly: The road from particles to colloidal molecules and crystals. *Angew. Chem. Int. Edit.* **50**, 360–388 (2011).
- [128] Boles, M. A., Engel, M. & Talapin, D. V. Self-assembly of colloidal nanocrystals: From intricate structures to functional materials. *Chem. Rev.* **116**, 11220–11289 (2016). PMID: 27552640.
- [129] Rechtsman, M. C., Stillinger, F. H. & Torquato, S. Optimized interactions for targeted self-assembly: Application to a honeycomb lattice. *Phys. Rev. Lett.* **95**, 228301 (2005).
- [130] Rechtsman, M., Stillinger, F. & Torquato, S. Designed interaction potentials via inverse methods for self-assembly. *Phys. Rev. E* **73**, 011406 (2006).
- [131] Torquato, S. Inverse optimization techniques for targeted self-assembly. *Soft Matter* **5**, 1157–1173 (2009).
- [132] Torquato, S. Optimal design of heterogeneous materials. *Annu. Rev. Mater. Res.* **40**, 101–129 (2010).
- [133] Jain, A., Bollinger, J. A. & Truskett, T. M. Inverse methods for material design. *AIChe J.* **60**, 2732–2740 (2014).
- [134] Sherman, Z. M., Howard, M. P., Lindquist, B. A., Jadrach, R. B. & Truskett, T. M. Inverse methods for design of soft materials. *J. Chem. Phys.* **152**, 140902 (2020).
- [135] Lindquist, B. A., Jadrach, R. B. & Truskett, T. M. Communication: Inverse design for self-assembly via on-the-fly optimization. *J. Chem. Phys.* **145**, 111101 (2016).
- [136] Tsai, C. L., Delaney, K. T. & Fredrickson, G. H. Genetic algorithm for discovery of globally stable phases in block copolymers. *Macromolecules* **49**, 6558–6567 (2016).
- [137] Paradiso, S. P., Delaney, K. T. & Fredrickson, G. H. Swarm intelligence platform for multiblock polymer inverse formulation design. *ACS Macro Letters* **5**, 972–976 (2016).
- [138] Khadiilkar, M. R., Paradiso, S., Delaney, K. T. & Fredrickson, G. H. Inverse design of bulk morphologies in multiblock

- polymers using particle swarm optimization. *Macromolecules* **50**, 6702–6709 (2017).
- [139] Khaira, G. S. *et al.* Evolutionary optimization of directed self-assembly of triblock copolymers on chemically patterned substrates. *ACS Macro Lett.* **3**, 747–752 (2014).
- [140] Kumar, R., Coli, G. M., Dijkstra, M. & Sastry, S. Inverse design of charged colloidal particle interactions for self assembly into specified crystal structures. *J. Chem. Phys.* **151**, 084109 (2019).
- [141] Miskin, M. Z., Khaira, G., de Pablo, J. J. & Jaeger, H. M. Turning statistical physics models into materials design engines. *Proc. Nat. Acad. Sci. USA* **113**, 34–39 (2016).
- [142] Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Assembly of nothing: equilibrium fluids with designed structured porosity. *Soft Matter* **12**, 2663–2667 (2016).
- [143] Piñeros, W. D., Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Inverse design of multicomponent assemblies. *J. Chem. Phys.* **148**, 104509 (2018).
- [144] Piñeros, W. D., Jadrlich, R. B. & Truskett, T. M. Design of two-dimensional particle assemblies using isotropic pair interactions with an attractive well. *AIP Adv.* **7**, 115307 (2017).
- [145] Adorf, C. S., Antonaglia, J., Dshemuchadse, J. & Glotzer, S. C. Inverse design of simple pair potentials for the self-assembly of complex structures. *J. Chem. Phys.* **149**, 204102 (2018).
- [146] Lindquist, B. A., Jadrlich, R. B. & Truskett, T. M. Communication: From close-packed to topologically close-packed: Formation of laves phases in moderately polydisperse hard-sphere mixtures. *J. Chem. Phys.* **148**, 191101 (2018).
- [147] Jaeger, H. M. & de Pablo, J. J. Perspective: Evolutionary design of granular media and block copolymer patterns. *APL Mater.* **4**, 053209 (2016).
- [148] Tokuda, S., Nagata, K. & Okada, M. Simultaneous estimation of noise variance and number of peaks in bayesian spectral deconvolution. *J. Phys. Soc. Japan* **86**, 024001 (2017).
- [149] Bär, M., Hegger, R. & Kantz, H. Fitting partial differential equations to space-time dynamics. *Phys. Rev. E* **59**, 337–342 (1999).
- [150] Müller, T. G. & Timmer, J. Parameter identification techniques for partial differential equations. *Int. J. Bifurcat. Chaos* **14**, 2053–2060 (2004).
- [151] Zhao, H., Storey, B. D., Braatz, R. D. & Bazant, M. Z. Learning the physics of pattern formation from images. *Phys. Rev. Lett.* **124**, 060201 (2020).
- [152] Voss, H. U., Timmer, J. & Kurths, J. Nonlinear dynamical system identification from uncertain and indirect measurements. *Int. J. Bifurcat. Chaos* **14**, 1905–1933 (2004).
- [153] Bongard, J. & Lipson, H. Automated reverse engineering of nonlinear dynamical systems. *Proc. Nat. Acad. Sci. USA* **104**, 9943–9948 (2007).
- [154] Schmidt, M. & Lipson, H. Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).
- [155] Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Nat. Acad. Sci. USA* **113**, 3932–3937 (2016).
- [156] Schaeffer, H. Learning partial differential equations via data discovery and sparse optimization. *Proc. Royal Soc. A* **473**, 20160446 (2017).
- [157] Rudy, S. H., Brunton, S. L., Proctor, J. L. & Kutz, J. N. Data-driven discovery of partial differential equations. *Sci. Adv.* **3**, e1602614 (2017).
- [158] Schaeffer, H. & McCalla, S. G. Sparse model selection via integral terms. *Phys. Rev. E* **96**, 023302 (2017).
- [159] Xu, H., Chang, H. & Zhang, D. Dl-pde: Deep-learning based data-driven discovery of partial differential equations from discrete and noisy data. *arXiv:1908.04463* (2019).
- [160] Bishop, C. M. *Pattern recognition and machine learning I*, vol. 1 (springer New York, 2006).
- [161] Kaipio, J. & Somersalo, E. *Statistical and computational inverse problems*, vol. 160 (Springer Science & Business Media, 2006).
- [162] Xun, X., Cao, J., Mallick, B., Maity, A. & Carroll, R. J. Parameter estimation of partial differential equation models. *J. Am. Stat. Assoc.* **108**, 1009–1020 (2013).
- [163] Campillo-Funollet, E., Venkataraman, C. & Madzvamuse, A. Bayesian parameter identification for turing systems on stationary and evolving domains. *Bull. Math. Biol.* **81**, 81–104 (2019).
- [164] Kazarnikov, A. & Haario, H. Statistical approach for parameter identification by turing patterns. *Journal of Theoretical Biology* **501**, 110319 (2020).
- [165] Zhao, H., Braatz, R. D. & Bazant, M. Z. Image inversion and uncertainty quantification for constitutive laws of pattern formation. *J. Comp. Phys.* **436**, 110279 (2021).
- [166] Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6**, 187–202 (2009).
- [167] Law, K., Stuart, A. & Zygalakis, K. *Data assimilation* (Springer, 2015).
- [168] Carrasi, A., Bocquet, M., Hannart, A. & Ghil, M. Estimating model evidence using data assimilation. *Q. J. R. Meteorol. Soc.* **143**, 866–880 (2017).
- [169] Evensen, G. *Data assimilation: the ensemble Kalman filter* (Springer, 2009).
- [170] Ito, S.-i., Nagao, H., Kurokawa, T., Kasuya, T. & Inoue, J. Bayesian inference of grain growth prediction via multi-phase-field models. *Phys. Rev. Mat.* **3**, 053404 (2019).
- [171] Yuan, Y. *et al.* Machine discovery of partial differential equations from spatiotemporal data. *arXiv:1909.06730* (2019).
- [172] Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005).
- [173] Tenorio, L. *An introduction to data analysis and uncertainty quantification for inverse problems*, vol. 3 (SIAM, 2017).

- [174] Varah, J. M. A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing* **3**, 28–46 (1982).
- [175] Ramsay, J. O., Hooker, G., Campbell, D. & Cao, J. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 741–796 (2007).
- [176] Rasmussen, C. E. *Gaussian processes in machine learning* (2003).
- [177] Graepel, T. Solving noisy linear operator equations by gaussian processes: Application to ordinary and partial differential equations. In *ICML*, 234–241 (2003).
- [178] Álvarez, M., Luengo, D. & Lawrence, N. D. Latent force models. vol. 5 of *Proceedings of Machine Learning Research*, 9–16 (PMLR, 2009).
- [179] Melkumyan, A. Operator induced multi-task gaussian processes for solving differential equations. In *Neural Information Processing Systems (NIPS) Workshop: New Directions in Multiple Kernel Learning* (2012).
- [180] Atkinson, S. & Zabaras, N. Structured bayesian gaussian process latent variable model: Applications to data-driven dimensionality reduction and high-dimensional inversion. *Journal of Computational Physics* **383**, 166 – 195 (2019).
- [181] Calderhead, B., Girolami, M. & Lawrence, N. D. Accelerating bayesian inference over nonlinear differential equations with gaussian processes. In Koller, D., Schuurmans, D., Bengio, Y. & Bottou, L. (eds.) *Advances in Neural Information Processing Systems 21*, 217–224 (2009).
- [182] Dondelinger, F., Husmeier, D., Rogers, S. & Filippone, M. Ode parameter inference using adaptive gradient matching with gaussian processes. vol. 31 of *Proceedings of Machine Learning Research*, 216–228 (PMLR, 2013).
- [183] Gorbach, N. S., Bauer, S. & Buhmann, J. M. Scalable variational inference for dynamical systems. In *Advances in Neural Information Processing Systems*, 4806–4815 (2017).
- [184] Raissi, M., Perdikaris, P. & Karniadakis, G. Numerical gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing* **40**, A172–A198 (2018).
- [185] Baake, E., Baake, M., Bock, H. G. & Briggs, K. M. Fitting ordinary differential equations to chaotic data. *Phys. Rev. A* **45**, 5524–5529 (1992).
- [186] Müller, T. & Timmer, J. Fitting parameters in partial differential equations from partially observed noisy data. *Physica D: Nonlinear Phenomena* **171**, 1–7 (2002).
- [187] Foucart, S. & Rauhut, H. *A Mathematical Introduction to Compressive Sensing* (Springer Science & Business Media, New York, NY, 2013).
- [188] Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. Royal Stat. Soc. B* **67**, 301–320 (2005).
- [189] Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
- [190] Shi, A.-C. Nature of anisotropic fluctuation modes in ordered systems. *J. Phys. Cond. Mat.* **11**, 10183 (1999).
- [191] Impéror-Clerc, M. Three-dimensional periodic complex structures in soft matter: investigation using scattering methods. *Interface focus* **2**, 589–601 (2012).
- [192] Bates, M. W. *et al.* Stability of the a15 phase in diblock copolymer melts. *Proc. Nat. Acad. Sci. USA* **116**, 13194–13199 (2019).